Mitochondrial DNA diversity of the Northern Bat (*Eptesicus nilssonii*) in the northern Palearctic

Anna Lempiäinen
Pro gradu – tutkielma
Tammikuu 2014
Perinnöllisyystiede
Biotieteiden laitos
Helsingin yliopisto

Tiedekunta – Fakultet – Faculty	Laitos – Institution– Department	
Biological and environmental sciences	Biosciences	
Tekijä – Författare – Author		
Anna Lempiäinen		
Työn nimi – Arbetets titel – Title		
Mitochondrial DNA diversity of the Northern Bat (*Eptesicus nilssonii*) in the northern Palearctic		
Oppiaine – Läroämne – Subject		
Genetics		
Työn laji – Arbetets art – Level	Aika – Datum – Month and year	Sivumäärä – Sidoantal – Number of pages
Master's thesis	January 2014	58 p. + supplements 14 p.

Tiivistelmä –Referat – Abstract

Several factors, such as the climate fluctuations during the Pleistocene ice age, have contributed to the geographical distribution of genetic variation in contemporary populations. Phylogeography studies the variation by connecting the genetic lineages of individuals with their geographical locations. One of the most popular markers used in these studies is mitochondrial DNA (mtDNA) due to its practical qualities. Mitochondrial DNA has revealed a whole new diversity of bat species and populations compared to conventional study methods using morphology alone. The study species, the Northern Bat (*Eptesicus nilssonii*), is a common, widely distributed and the most northern one of the Palearctic bat species, but its entire range has not yet been studied genetically. My aim is to study the mtDNA diversity of the Northern Bat mainly in the area of Fennoscandia and Latvia and compare this data with earlier published sequences from individuals located elsewhere in the northern Palearctic to elucidate its population structure and history.

The mtDNA diversity of the Northern Bat was examined from 146 individuals, of which mtDNA was sequenced of two different mtDNA markers, cytochrome *b* and control region. The DNA was obtained from tissue material of live bats and museum specimens. Additional data comprised 6 Northern Bat and 13 Serotine sequences. Sequences of a sister species, the Serotine (*E. serotinus*), were used as an outgroup. Three geographically embedded data sets from the northern Palearctic, northern Europe and Finland, were formed for examining population structure in different geographical scales.

The Northern Bat population was observed to be divided into two mitochondrial lineages; one located mostly in West (European lineage) and the other in East (Siberian lineage). These lineages seem to have diverged 0.85–1.1 million years ago based on a corrected cytochrome *b* distance of 1.7–2.2 %. On the control region, the European lineage showed considerable genetic diversity (π = 0.019, h = 0.966), and the Siberian lineage high haplotype diversity (h = 0.978) but relatively low nucleotide diversity (π = 0.009). The European lineage was further divided into four genetically different groups, clusters, but the Siberian lineage formed only a single cluster. The variation were geographically structured on the north European scale (Φ_{ST} = 0.07), but not on the Finnish scale (Φ_{ST} = -0.002). One of the Finnish colonies was found to comprise both lineages, while other sampled colonies consisted of only individuals of the European lineage.

The European and Siberian lineages have presumably diverged in isolation through the Middle Pleistocene times, and, despite the genetic distinction between the lineages, they most likely represent interbreeding conspecific populations. Thus, the species' nomenclature needs no changes. At the north European scale, the high nucleotide diversity observed in the European lineage resulted from the four clusters, which most likely formed in the major refugia in Europe, whereas the low nucleotide diversity in the Siberian lineage resulted from the single cluster, which presumably originated in a single refugium in Central Asia. On the Finnish population scale, both lineages were observed, but the population in general was not geographically structured. The colony comprising both lineages was most likely an indication of the lineages interbreeding. Fennoscandia was evidently recolonized by the European lineage via two routes (Denmark/Sweden and the Baltic countries), and the Siberian lineage via one (Karelia) from Russia. The lineages have come into contact for a second time in Finland and presumably in the European Russia after the end of the last glaciation. Thus, the Pleistocene ice age was a substantial contributor to the observed contemporary population structure of the Northern Bat.

Avainsanat – Nyckelord – Keywords

Phylogeography, *Eptesicus nilssonii*, cytochrome *b*, control region, population history, glacial refugia

Ohjaaja tai ohjaajat – Handledare – Supervisor or supervisors

Risto Väinölä and Eeva-Maria Kyheröinen

Säilytyspaikka – Förvaringställe – Where deposited

Library of the Department of Biosciences, Viikki campus library

Muita tietoja – Övriga uppgifter – Additional information

Tiedekunta – Fakultet – Faculty	Laitos – Institution– Department	
Bio- ja ympäristötieteellinen	Biotieteiden	
Tekijä – Författare – Author		
Anna Lempiäinen		
Työn nimi –Arbetets titel – Title		
Pohjanlepakon (*Eptesicus nilssonii*) mitokondrio-DNA:n diversiteetti palearktisen alueen pohjoisosassa		
Oppiaine – Läroämne – Subject		
Perinnöllisyystiede		
Työn laji – Arbetets art – Level	Aika – Datum – Month and year	Sivumäärä – Sidoantal – Number of pages
Pro gradu –tutkielma	Tammikuu 2014	58 s. + liitteet 14 s.

Tiivistelmä –Referat – Abstract

Monet tekijät, kuten pleistoseenin jääkauden aikaiset ilmaston vaihtelut, ovat vaikuttaneet geneettisen variaation maantieteelliseen jakaumaan eliöillä. Fylogeografia tutkii tätä variaatiota yhdistämällä yksilöiden geneettiset linjat niiden maantieteellisiin sijainteihin. Yksi suosituimmista geneettisistä markkereista, joita tällaisissa tutkimuksissa on käytetty, on mitokondrio-DNA (mtDNA) sen käytännöllisten ominaisuuksien takia. Mitokondriaalinen DNA on paljastanut kokonaan uudenlaisen lepakkolajien ja – populaatioiden variaation verrattuna perinteisiin morfologiaan perustuviin tutkimustapoihin. Tutkimuslaji, pohjanlepakko (*Eptesicus nilssonii*), on yleinen, laajalle levinnyt ja kaikista palearktisen lepakkolajeista pohjoisin, mutta sitä ei ole tutkittu geneettisesti vielä koko esiintyvyysalueeltaan. Tavoitteenani on tutkia pohjanlepakon mtDNA:n diversiteettiä pääosin Fennoscandian ja Latvian alueilta ja verrata tätä aineistoa aikaisemmin julkaistuihin sekvensseihin yksilöistä, jotka sijaitsevat muualla pohjoisen palearktisen alueella selvittääkseni pohjanlepakon populaatiorakennetta ja historiaa.

Pohjanlepakon mtDNA:n diversiteettiä tutkittiin 146 yksilöstä, joiden mtDNA sekvensoitiin kahdesta eri markkerista, sytokromi *b*:stä ja kontrollialueesta. DNA saatiin elävien lepakoiden sekä museonäytteiden kudosmateriaalista. Lisäaineistossa oli 6 pohjanlepakko- ja 13 etelänlepakkosekvenssiä. Sisarlajin, etelänlepakon (*E. serotinus*), sekvenssejä käytettiin ulkoryhmänä. Kolme sisäkkäistä maantieteellistä aineistoryhmää pohjoiselta palearktiselta, pohjoisesta Euroopasta ja Suomesta muodostettiin, jotta populaatiorakennetta voitiin tarkastella erilaisissa maantieteellisissä tasoissa.

Pohjanlepakkopopulaation havaittiin jakautuvan kahteen mitokondriolinjaan; yksi esiintyi pääasiassa lännessä (Euroopan linja) ja toinen idässä (Siperian linja). Nämä linjat eriytyivät luultavasti 0,85–1,1 miljoonaa vuotta sitten korjatun sytokromi *b*:n etäisyyden 1,7–2,2 % perusteella. Kontrollialueella Euroopan linjassa oli korkea geneettinen diversiteetti (π = 0,019, h = 0,966), ja Siperian linjassa korkea haplotyyppidiversiteetti (h = 0,978) mutta suhteellisen matala nukleotididiversiteetti (π = 0,009). Euroopan linja jakaantui edelleen neljään geneettisesti erilaiseen ryhmään, klusteriin, mutta Siperian linja muodosti vain yhden klusterin. Näillä klustereilla havaittiin maantieteellistä rakennetta pohjoisen Euroopan tasolla (Φ_{ST} = 0,07), mutta ei Suomen tasolla (Φ_{ST} = -0,002). Yhden suomalaisista kolonioista havaittiin koostuvan molempien linjojen yksilöistä, kun taas muissa tutkituissa kolonioissa oli vain Euroopan linjan yksilöitä.

Euroopan ja Siperian linjat luultavasti muodostuivat eristäytyneinä toisistaan Keski-pleistoseenin aikana ja, huolimatta linjojen välisestä geneettisestä eroavaisuudesta, linjat ovat todennäköisimmin keskenään lisääntyviä samaan lajiin kuuluvia populaatioita. Siten lajinimeen ei tarvita muutoksia. Pohjoisen Euroopan tasolla tarkasteltuna Euroopan linjassa havaittu korkea nukleotididiversiteetti oli seurausta linjan neljästä ryhmästä, jotka olivat todennäköisesti syntyneet Euroopan merkittävissä refugioissa, kun taas Siperian linjan matala nukleotididiversiteetti oli seurausta linjan muodostamasta ainoasta ryhmästä, joka luultavasti muodostui yhdessä refugiossa Keski-Aasiassa. Suomen populaation tasolla molemmat linjat havaittiin, mutta yleisesti maantieteellistä populaatiorakennetta ei ollut. Kolonia, josta löydettiin kumpaakin linjaa, oli todennäköisimmin osoitus linjojen risteytymisestä keskenään. Euroopan linja kolonisoi Fennoskandian luultavasti kahta reittiä pitkin (Tanska/Ruotsi ja Baltian maat) ja Siperian linja yhtä reittiä pitkin (Karjala) Venäjältä. Linjat kohtaavat toisensa uudelleen Suomessa ja luultavasti eurooppalaisen Venäjän alueella viimeisen jäätiköitymisen päätyttyä. Täten Pleistoseenin jääkausi on ollut merkittävä tekijä pohjanlepakolla havaitun nykyisen populaatiorakenteen muodostumiselle.

Avainsanat – Nyckelord – Keywords
Fylogeografia, *Eptesicus nilssonii*, sytokromi *b*, kontrollialue, populaatiohistoria, glasiaalirefugio
Ohjaaja tai ohjaajat – Handledare – Supervisor or supervisors
Risto Väinölä ja Eeva-Maria Kyheröinen
Säilytyspaikka – Förvaringställe – Where deposited
Biotieteiden laitoskirjasto, Viikin kampuskirjasto
Muita tietoja – Övriga uppgifter – Additional information

CONTENTS

ABBREVIATIONS

AMOVA	Analysis of Molecular Variance
COI	cytochrome oxidase subunit I
cyt *b*	cytochrome *b*
F_{ST}	fixation index
h	haplotype diversity
HSP	heavy-strand promoter
HVI, HVII, HVIII	Hypervariable segments I, II and III
JC distance	Jukes-Cantor distance
kb	kilobase
LSP	light-strand promoter
mtDNA	mitochondrial DNA
Mya	million years ago
Myr	million years
ND1	NADH dehydrogenase subunit I
NJ	neighbor-joining
oriH	heavy-strand origin of replication
UV	ultraviolet
Φ_{ST}	fixation index
π	nucleotide diversity
π_T	total nucleotide diversity

1 INTRODUCTION

1.1 Phylogeography

Phylogeography is a field of research that combines disciplines of the macro- and microevolutionary levels. These disciplines include for instance phylogeny, i.e. evolutionary history and population genetics, of which the first focuses on major trends among higher taxa, and the latter traditionally addresses the intraspecific geographical population structure and population history. Phylogeography aims to connect genealogy, or genetic lineages, with geography to study their relationships and to reveal the processes that have caused the observed distribution of genetic variation in contemporary populations (Avise *et al.* 1987). In animal populations, the genetic structure can vary from total uniformity across geographic distributions (e.g. Kvist *et al.* 1999) to intense structure (e.g. Rodríguez *et al*. 2013) subdivided by e.g. geographical and ecological boundaries or behavioral consequences (Avise 2004). The genetic structure of populations is usually influenced by past events, and their occurrence in time can be determined by genetic data collected from the populations (Hamilton 2009).

Mitochondrial DNA (mtDNA) is the most used molecular marker in phylogeographic studies (Randi 2000), in which also other highly variable markers of the nuclear genome are often included for additional information. Only during the past couple of decades phylogeographic studies have revealed vast amounts of new information on plant and animal biogeographical histories and population structure (Avise 2000), including the fascinating time frame of the Quaternary period, comprising the Pleistocene ice ages.

1.2 Pleistocene ice ages and their effect on northern Palearctic biota

The genetic diversity of many natural populations has been shaped by the Quaternary ice ages (Hewitt 1996). During the Pleistocene epoch, which began about 2.5 million years ago (Mya), the climate changed repeatedly from cold glacial to warm interglacial conditions and induced continental ice sheets to grow and shrink. The temperate northern hemisphere was affected most severely by these climatic fluctuations (Webb & Bartlein 1992). The temperate biota contracted and expanded with the changing climate, and during the severe glacial stages plants and animals were

forced to move south and settle to the remaining habitable sanctuaries, i.e. refugia (Hewitt 1996, 1999), which led to genetic differentiation in the allopatric populations (Bilton *et al*. 1998). Iberia, Apennines (Italy) and the Balkans have been accepted as the main refugial areas (e.g. Hewitt 1999), but central Europe and Western Asia have also been suggested as sources of the present central and north European populations (Bilton *et al*. 1998; Provan & Bennett 2008). The three main refugia are distinct due to the geographical traits of Europe: it is a large east-west oriented peninsula, and the Mediterranean Sea forms a strong barrier of dispersal in the South. Also the east-west oriented mountain ranges of the Alps and the Pyrenees impeded the northward movement of the biota from the Apennines and Iberia during the warm interglacials, thus enhancing the isolation of the populations (Taberlet *et al*. 1998).

Approximately 18 000 years ago, at the end of the Late Pleistocene, the ice sheets began to recede and the global climate became warmer. The biota migrated northwards following their optimal circumstances (Huntley & Webb 1989). This expansion of the refugial populations has been associated with an observed decreasing south-to-north trend of genetic variation in some species. The trend was caused by a series of bottlenecks when the biota spread from the leading edge of the refugial population, and led to a loss of alleles and decreasing genetic diversity (Hewitt 1996).

In some of the studied terrestrial European and Scandinavian vertebrates, the intraspecific genealogical lineages, which formed in separate refugia, were found to have come to secondary contact in the Fennoscandian area, for instance among the following species: Moor Frog (*Rana arvalis*: Knopp & Merilä 2009), Bank Vole (*Myodes glareolus*; Tegelström 1987), Field Vole (*Microtus agrestis*; Jaarola & Tegelström 1995) and Brown Bear (*Ursus arctos*; Taberlet *et al*. 1995). These areas of considerable genetic diversity, which has been caused by the secondary contact zones, have also been located in central Europe (Taberlet *et al*. 1998) and even in Asia at least among passerine birds (Aliabadian *et al*. 2005). Due to the relative recentness of the Holocene recolonization and the Pleistocene population isolations, fast evolving mitochondrial DNA is used in most studies that cover the Quaternary because of its fine resolution (Hewitt 1999).

1.3 Molecular evolution of the mammalian and bat mitochondrial DNA

The mammalian mitochondrial genome is a small, circular approximately 16 kilobase (kb) sized DNA molecule, which is maternally inherited in most species (Rabinowitz & Swift 1970; Hutchison *et al*. 1974). It consists of approximately 37 genes, of which 22 code for transfer RNAs,

two ribosomal RNAs and 13 protein-coding messenger RNAs that code for proteins mainly involved in the electron transport and oxidative phosphorylation (Figure 1.1). The mitochondrial genome is very efficiently arranged; it lacks introns, the spacers between genes are small or lacking and some genes can overlap by a few base pairs, although, duplications or insertions can be found in the control region (Avise 2004; Meganathan *et al*. 2012).

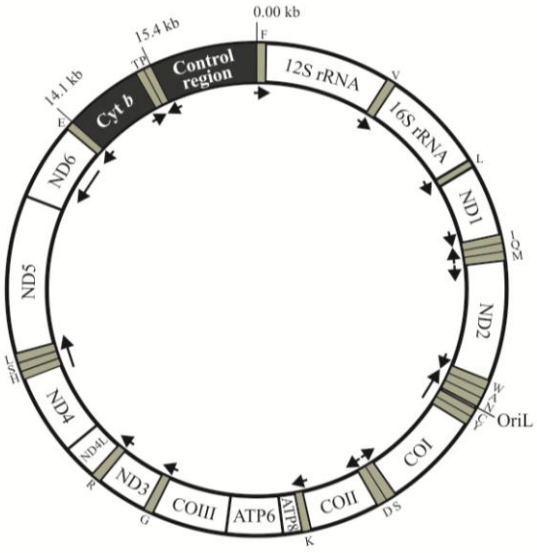


Figure 1.1 Organization of the bat mitochondrial genome after Meganathan *et al*. (2012). Arrows pointing counter clockwise depict genes transcribed from the L-strand, and arrows pointing clockwise are genes transcribed from the H-strand. The two mtDNA regions used in this study are marked with dark grey.

Mitochondrial DNA has many useful qualities that can be utilized in studies of evolutionary history on intra- and interspecies levels (Ballard & Whitlock 2004). The haploid mtDNA passes on predominantly through the female lineage (Birky 2005), and because it rarely recombines, the mitochondrial genome can be seen as a single genealogical unit (Avise 1989; Birky 2001). The number of mitochondria is generally high (~800) in most cells and each mitochondrion can contain multiple copies of the mitochondrial genome (Scheffler 1999), which usually are all identical within an individual, but vary between individuals at the intraspecific level (Avise 1986). The ample copy number is valuable when the samples in question are old or in poor condition, since then at least some mtDNA that is usable for the PCR amplification will have survived even when the nuclear DNA itself has perished (Pääbo *et al.* 2004). Different regions of mtDNA have different evolutionary rates (Pesole *et al*.1999), which can be used in studies on diverse taxonomic levels and time scales. The regions evolving slower are more suitable for interspecies studies and those evolving faster for population level studies (Avise 2004).

In bat studies, mitochondrial markers, which are often used along with additional nuclear markers, have frequently discovered cryptic species – two or more species within a same nominal species –, e.g. for the genera of *Pipistrellus*, *Hypsygo*, *Plecotus*, *Eptesicus* and *Myotis* in western Palearctic (Mayer *et al*. 2007) and for the *Miniopterus* genus in Madagascar (Goodman *et al*. 2009). Thus, when using solely morphology or behavior-based methods distinct bat species can be left unrecognized.

Cases of non-monophyly of the mitochondrial lineage have also been frequently found in various bat studies (Hoffman *et al*. 2003; Berthier *et al*. 2006; Mayer *et al*. 2007; Mao *et al*. 2010). These indicate that hybridization and introgression, i.e. gene leakage from one species to another, between diverged lineages or species has taken place during the history of the study taxon. In such cases, the molecular history is different from the species history, which could be falsely inferred if only mtDNA is used (Machado & Hey 2003).

At a higher systematic level, comparing mitochondrial genomes suggested that the microbat suborder Microchiroptera is not monophyletic (Teeling *et al*. 2000), which later led to the reorganization of the order of Chiroptera (Hutcheon & Kirsch 2006). The mitochondrial genome of bats has the usual mammalian gene organization (Figure 1.1). Out of the over 1200 bat species, the complete DNA sequence of the mitochondrial genome is known from less than 20 species (Meganathan *et al*. 2012). However, sequences of some individual segments of the mitochondrial genome, such as the control region, NADH dehydrogenase subunit 1 (ND1) and cytochrome *b*, have been researched widely. In this study, segments of two popular mtDNA markers were utilized: cytochrome *b* and the control region.

1.3.1 Cytochrome *b*

In the mitochondrial genome, cytochrome *b* is a 1.4 kb long gene, which produces a protein that operates as a part of the electron transport chain in the inner membranes of the mitochondria. It is embedded in the inner membrane by its eight transmembrane helices (Figure 1.2). Cytochrome *b* is the only cytochrome coded by mitochondrial DNA (Esposti *et al*. 1993).

The structure and function of the cytochrome *b* protein limit its evolution at the DNA level, since non-synonymous substitutions change the biochemical qualities of the protein product. Although the rate of evolution is slow when the non-synonymous codon sites are considered, evolving is

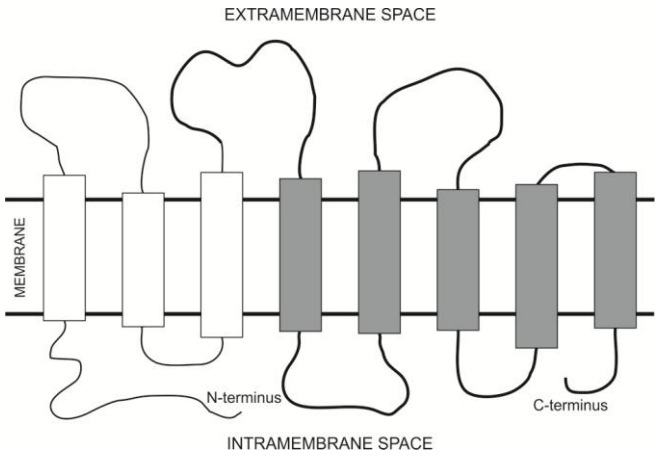


Figure 1.2 Structure of cytochrome *b* (after Howell 1989). The shaded and bolded parts of the protein correspond to the section of the gene used in this study.

relatively fast at the synonymous sites (Irwin *et al*. 1991). The conserved non-synonymous substitutions can be utilized for studies of deeper phylogenetic relationships, whereas the variable synonymous sites can be used for population level questions (Meyer 1994).

1.3.2 Control region

The control region is the only large, approximately 1 kb long section of non-coding DNA in mammalian mitochondrial DNA. It functions as an origin of replication for the heavy-strand and contains promoters for transcription of both light- and heavy-strands (Wilkinson *et al*. 1997; Avise 2004). Despite its role in replication and transcription, substitutions accumulate into the control region sequence relatively fast (Sun *et al*. 2009). The control region is generally divided into three main domains. The central domain is the most conserved in evolution, and each domain consists of alternating variable and conserved sequence blocks (Figure 1.3; Wilkinson *et al*. 1997; Pesole *et al*. 1999; Sun *et al*. 2009).

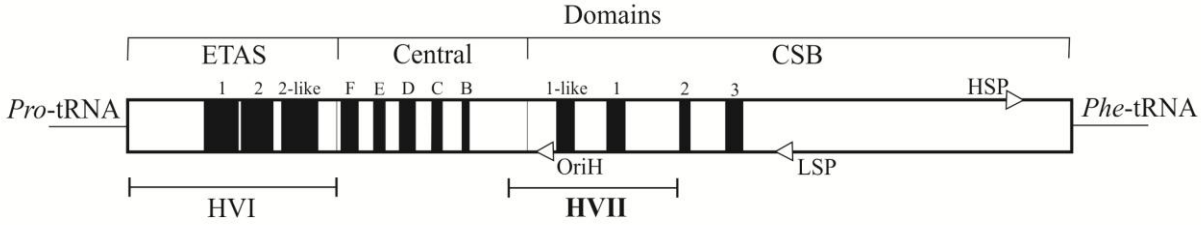


Figure 1.3 Structure of the control region (after Sun *et al*. 2009). The approximate locations of HVI and HVII are estimated from the information of human mitochondrial hypervariable segments (Lutz *et al*. 2000). Conserved sequence blocks (Sun *et al*. 2009) are marked with black bars. OriH is the origin of replication of the heavy-strand; LSP and HSP are the light- and heavy-strand promoters (after Wilkinson *et al*. 1997).

Hypervariable (HV) segments I and II are the most used control region markers in bat studies; sometimes the complete control region is used (e.g. Xu *et al*. 2010). In humans and birds HVIII has also been recognized and used for distinction of different individuals (Lutz *et al*. 2000; Crochet & Desmarais 2000).

1.3.3 Applications for mtDNA data: the molecular clock and mismatch distribution

The mitochondrial DNA sequence data can be used to examine genetic changes in past populations that have left characteristic traces observable in the DNA of present populations. One of the methods for studying these processes is the molecular clock, which enables dating the events of divergence of taxa or lineages in sidereal time (Hamilton 2009). The idea of the molecular clock was introduced by Zuckerkandl & Pauling (1965). It assumes that the substitution rate, i.e. the amount of divergence, is roughly constant through time and is thus proportional to the time of separation of the taxa. The substitutions measured must be neutral to natural selection so that constant mutation rates can be at least roughly assumed (Hamilton 2009). Thus, the hypothesis of the molecular clock is based on the neutral theory, which was developed by Kimura (1968).

One of the essential features of the molecular clock is that the estimation of the divergence times can be accomplished without comprehensive fossil records, assuming that substitution rates are similar among different taxa (Arbogast *et al*. 2002). However, a clock calibration based on fossil records is needed at least initially to obtain more accurate estimates, but interpreting comparisons of DNA sequence data against fossils can be challenging, since the DNA sequence data might overestimate and the fossil data underestimate the divergence times, and some species cannot be distinguished from one another merely by the morphology of the skeleton (Avise 2004; Hamilton 2009). Calibration can also be based on biogeographical events, such as formation of an island (Fleischer *et al*. 1998; Bromham & Penny 2003), but still the interpretation is not always straightforward (Avise *et al*. 1992). When comparing more distant taxa, substitution rates can vary considerably between them, in which case the molecular clock cannot be considered as a universal timepiece. The divergence of closely related species can be estimated by a taxonomically local clock, since it is probable that their rates of molecular evolution are similar (Arbogast *et al*. 2002).

By applying the molecular clock, it is also possible to estimate the dates of past demographic events or historical changes in population size from extant populations. A simple method for examining events in population history is the distribution of all individual pairwise comparisons within a

population. This method is called mismatch distribution. Different kinds of events in population history are expected to produce distinct shapes of mismatch distributions, such as the theoretical L-shaped curve caused by a significant reduction in population size (Rogers & Harpending 1992). As an alternative to mismatch distributions, demographic signals are often studied by coalescence simulations of the mitochondrial genealogy (Drummond *et al*. 2002). In coalescence two or more lineages converge to a single common ancestral lineage when the lineages are traced back in time (Hamilton 2009).

1.5 The Northern Bat

The Northern Bat (*Eptesicus nilssonii*) is a member of the family Vespertilionidae, and is one of the 23 different species belonging to the genus *Eptesicus*. The range of the mostly non-migrating Northern Bat covers Fennoscandia, part of the continental Europe and southern Russia and extends to the Pacific coast and Northern Japan, Hokkaido (Figure 2.1; Wilson & Reeder 2005). The range the Northern Bat inhabits is mostly temperate and boreal extending to the mountain woodlands in the northern Palearctic (Stubbe *et al*. 2008). Northern Bats have adapted to the harsh conditions of their boreal environment (Rydell 1989), and can thrive in different habitats ranging from close proximity of streams and lakes to open grasslands, forests and even to urban areas (Dietz *et al.* 2009).

In Finland, the Northern Bat is the most numerous bat species (Siivonen & Wermundsen 2008). Thus, it is not considered endangered (Liukko *et al*. 2010) and is classified as a Least Concern species in the IUCN classification. Nonetheless, as well as all the other bat species in Finland, Northern Bat is protected by national legislations, as in most other countries in the range of this species. Other conservation measures include the Bonn Convention (Eurobats), the Annex IV of EU Habitats and Species Directive and Natura 2000, which protects some of the Northern Bat habitats (Stubbe *et al*. 2008).

Previously the Northern Bat has been examined among other species of the genus *Eptesicus* (Juste *et al*. 2013; Artyushin *et al*. 2009) or among other vespertilionid bat genera (e.g. Mayer & von Helversen 2001; Ibáñez *et al*. 2006; Kruskop *et al*. 2012) using mitochondrial and nuclear markers. These studies have revealed some very exciting information of the Northern Bat. Interbreeding of the Northern Bat and its sister species, the Serotine (*E. serotinus*), was first suggested by Mayer and von Helversen (2001) and later reaffirmed by Juste *et al*. (2013). The latter authors came to the

conclusion that the Northern Bat mtDNA was introgressed to the Serotine. Introgression has also been detected in some other bat species (in genera *Uroderma* (Tent-Making Bats); Hoffman *et al*. 2003, *Pteropus* (Flying Foxes); Webb & Tidemann 1995, *Rhinolophus* (Horseshoe Bats); Mao *et al*. 2013, *Scotophilus* (Yellow/House Bats); Vallo *et al*. 2012 and *Myotis* (Mouse-Eared Bats); Berthier *et al*. 2006), but the amount will presumably rise with further research given that every fourth mammal species is a bat (Berthier *et al*. 2006) and many of them have not been studied by molecular markers or at all. Furthermore, Artyushin *et al*. (2009) discovered that the Northern Bat comprised two distinct mitochondrial lineages that the authors named by the geographical locations from which the lineages were found. The European lineage was detected only in Europe; west from the proximity of the Russian border, and the Siberian lineage was detected only in Russia. Later studies (Kruskop *et al*. 2012; Juste *et al*. 2013) have confirmed this intraspecific geographical division. However, all of these larger scale studies that were accomplished in the European and Russian areas disregarded Fennoscandia. Smaller scale studies of refugia and recolonization pathways have been conducted for some of those northwest Palearctic bat species, of which distribution stretches at least to the southern parts of the Nordic countries (e.g. Petit *et al*. 1999; Juste *et al*. 2004; Ibáñez *et al*. 2006), but also these studies ignored Fennoscandian bat populations.

1.6 Aims of the study

The general aim of my study is to provide basic knowledge on the genetic diversity of the Northern Bat (*Eptesicus nilssonii*) in Fennoscandia and to use it for clarifying its population structure and history.

More specifically, I first used mitochondrial DNA sequences of cytochrome *b* gene to set a phylogenetic framework for the broad scale northern Palearctic data of the Northern Bat. This data serves as a basis for more detailed population level studies.

After this I examined genetic variation and population structure of the Northern Bat in different geographical levels by using mitochondrial control region DNA sequences of broad scale north European and more local Finnish data.

The following are the central questions of this study:

1) Are the northern Palearctic and Finnish Northern Bat populations geographically structured?

2) Does the Fennoscandian Northern Bat population consist either both of the intraspecific mitochondrial lineages, i.e. the European lineage and Siberian lineage, or of only one and if so, of which lineage is it?

3) When did the Northern Bat lineages diverge?

4) Are there any signs of ancient population division into a single refugium or multiple refugia during glaciations in the present Northern Bat population? Which post-glacial colonization routes might Northern Bats have migrated along?

5) Can two segments of the mitochondrial genome with different mutation rates be compared with each other? Are these regions suitable for studies of systematics and population genetics?

2 MATERIALS AND METHODS

2.1 Study area and samples

In this study three embedded geographical scales were selected for sampling: A large scale area of the Northern Palearctic, a mid-sized area of North Europe (Fennoscandia and Latvia), and a local scale including only Finland, which is also a part of the larger scales (Figure 2.1). The purpose was to examine possible population structures of the Northern Bat in the different geographical scopes.

Starting at the local scale, the base for the Finnish data was obtained from museum samples from the Zoological Museum of the Finnish Museum of Natural History (MZH) in Helsinki and from the Zoological Museum of the University of Oulu (ZMUO). The museum samples were older tissue preserved by freezing or alcohol. The Finnish data was supplemented with fresh tissue samples of live bats collected from various locations in Southern Finland during the summers of 2010 and 2012 (Figure 2.1; Appendix A). The total Finnish data consisted of 102 individuals.

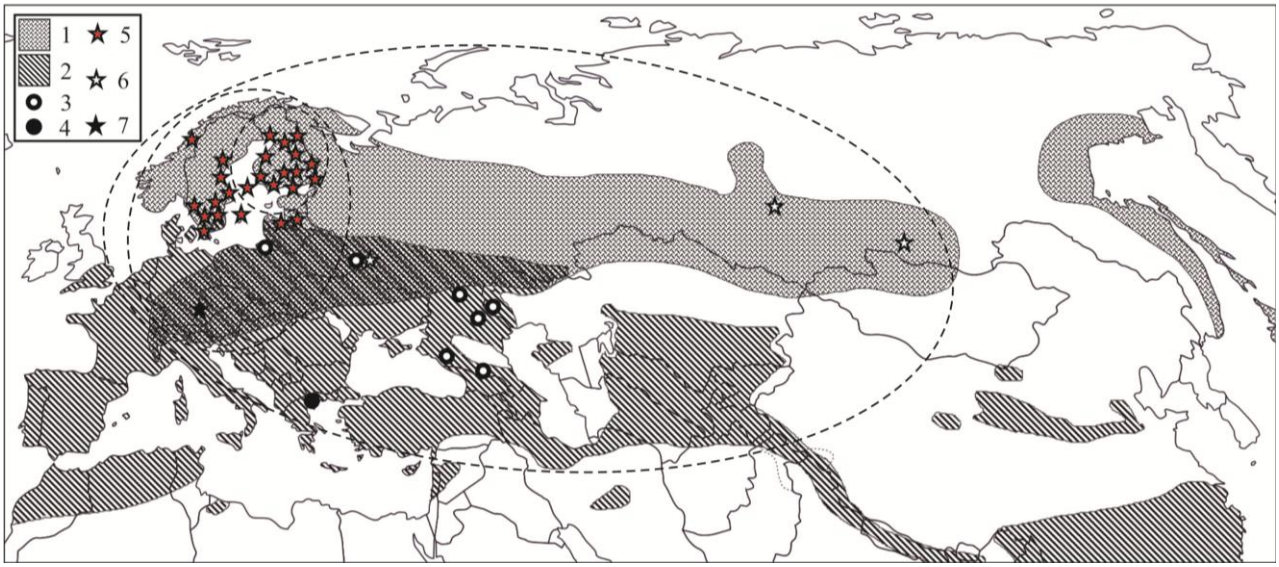


Figure 2.1 Sampling locations and distributions of the Northern Bat (*Eptesicus nilssonii*) and Serotine (*E. serotinus*), which is used as an outgroup species (ranges redrawn for both species after http://maps.iucnredlist.org range maps). The dashed line circles illustrate the three embedded geographical scales of this study: Northern Palearctic, North Europe and Finland. 1—Range of the Northern Bat, 2—Range of the Serotine, 3—Sample locations of Serotine (Artyushin *et al*. 2009), 4— Serotine (Ruedi & Mayer 2001), 5—Northern Bat (this study), 6—Northern Bat (Artyushin *et al*. 2009), 7—Northern Bat (Ruedi & Mayer 2001). Easternmost Eurasia and Japan are also inhabited by Northern Bats, but the area was excluded from this study to confine the thesis.

For the North European scale, museum samples were obtained from the Department of Vertebrate Zoology of the Swedish Museum of Natural History (NRM), from Helgeland museum (RMZ) in Norway and from the Latvian University of Agriculture (LUA): a total of 44 individuals were

sampled. The collection dates of all of the museums sampled bats varied from 1913 to 2010. MSc Eeva-Maria Kyheröinen, a licensed bat ringer, was responsible for the acquisition of both the museum and fresh tissue materials. A list of the Finnish and foreign samples can be found in Appendix A.

The data of the northern Palearctic and North European scales were supplemented by 6 Northern Bat (*Eptesicus nilssonii*) sequences and 13 sequences of an outgroup species, the Serotine (*E. serotinus*; outgroup is "a taxon phylogenetically outside the clade of interest", Avise 2004). These cytochrome *b* sequences were obtained from previous studies (Ruedi & Mayer 2001; Artyushin *et al*. 2009), and were retrieved from GenBank (see list in Appendix B for GenBank Accession Numbers).

2.2 Trapping methods of live bats

Bats were mainly trapped by a harp trap, which was first depicted by Constantine (1958), and later Tuttle (1974) introduced some improvements to it. The harp trap consists of two parallel 2.4 × 1.8 m aluminium tube frames and many vertical monofilament fishing lines stretched to each frame every 2.5 centimeters (Figure 2.2). A white canvas pouch, partly covered with polyethylene, is located under the lines. The working principle of the harp trap is that a flying bat avoids the first set of strings, but collides with the next one and falls into the pouch. The bats are prevented from escaping by a slippery polyethylene covering that extends into the pouch. It also protects the pouch from mild rain.

The bat trapping success of the harp trap was enhanced by the aid of an acoustic lure, which is a device that plays species specific high frequency social calls (Hill & Greenaway 2005). The loudspeaker was attached to the harp trap frame and the battery, connected to the loudspeaker by a long cord, was placed on the ground. An ultrasound detector had to be used to check that the siren was on and that it played the sound of the correct species, since the ultrasonic bat calls cannot be heard by human ear.

In addition, at a few sites some bats were caught by a hand net and mist nets, which are also used for trapping birds. Mist nets were mounted close to the harp trap between two fishing rods that were anchored to the ground with tent pegs.

Figure 2.2 A harp trap for trapping bats and tissue sampling from a live bat. On the left, the small black box attached to the closer vertical aluminium tube is the loudspeaker of the acoustic lure and the power source is located inside the clear box on the ground between the legs of the trap. On the right, the bat is Daubenton's bat (*Myotis daubentonii*).

2.3 Tissue sampling

Museum specimens were sampled for tissue either from the thorax near the wing, if there was enough muscle tissue, or from the wing membrane, if the muscle tissue had been degraded. I performed the sampling of museum samples only on the specimens that were obtained from the Zoological Museum of the Finnish Museum of Natural History (MZH). The sampling was performed in the laboratory of the Zoological Museum. Other museums' specimens were sampled by the staff in each museum, except specimens from the Zoological Museum of the University of Oulu (ZMUO) were sampled by MSc Eeva-Maria Kyheröinen. Altogether, 111 tissue samples were obtained from the museum specimens. Live bats were sampled by punches of 3 mm across taken from the wing between the 4th and 5th finger (Figure 2.2). The sampling was done with a specifically for this purpose designed tool that collects a round piece of wing tissue from the bat. The small pieces of tissue were placed in 96 % ethanol and stored in a refrigerator until further use (Worthington Wilmer & Barratt 1996). Fresh tissue samples were obtained from 40 Northern Bats.

2.4 DNA extraction and PCR amplification of mtDNA

All DNA work was conducted in the laboratory of the Finnish Museum of Natural History by me unless told otherwise. Bat DNA was extracted using a commercial Nucleospin® Tissue Protocol extraction kit following the instructions of Standard protocol for human or animal tissue and cultured cells with minor modifications (Appendix C). In preliminary experiments (by museum technician Katja Nylund) this method was compared with a simple salt extraction protocol, in terms of testing if the salt protocol was more efficient in yielding DNA than the kit. Most samples extracted by the salt protocol resulted with stronger bands to some extent when examined on agarose gel, but the difference was not considered significant. The kit method was also quicker and easier to conduct, thus it was chosen for this study. Nevertheless, a few salt extracted sequences (Fin4, Fin9–Fin11; listed in Appendix A) were included into the data of this study, since they were ready for further use and of good quality.

Two different segments of bat mitochondrial DNA were amplified by PCR: The first half of the coding cytochrome *b* gene and the non-coding hypervariable segment II (HVII) of the control region. Most primers were obtained from earlier bat studies, but one new primer was designed by senior curator Risto Väinölä specifically for this study (Table 2.1).

Table 2.1 Primers used for amplification of the partial cytochrome *b* (cyt *b*) gene and HVII sequences.

Primer	Target	Primer sequence (5'- 3')	Reference
EptL_486	cyt *b*	ATGAATTTGAGGAAGCTTTTCCGTAGA	Artyushin *et al*. 2009
H15395_pip	cyt *b*	CAGCTTTGGGTGTTGATGGTGG	Artyushin *et al*. 2009
ER63	HVII	CATCTGGTTCTTACTTCAGG	Petit *et al*. 1999
ER88	HVII	CAGCTTTGGGTGTTGATGGTGG	Petit *et al*. 1999
ER intH	HVII	TGCATATGTCCTGCGACCATG	this study

Samples were amplified in a 20 µl volume PCR reaction mix composed of 2 µl 10 × $(NH_4)_2SO_4$ *Taq* buffer (Fermentas), 1.2 µl $MgCl_2$ (Fermentas), 0.2 µl of *Taq* DNA polymerase (0.04 units (U); Fermentas), 0.4 µl dNTPs, 1 µl each primer [10 pMol/µl], 12.2µl water and 2 µl of the 50 µl DNA extract (average DNA concentration of 17 ng).

The PCR program begins with an initial denaturation step at 95°C for 3 minutes and the following steps were repeated 40 times: denaturation at 95°C for 45 seconds, annealing at 54°C for 45 seconds

and extension in 72°C for 1 minute. A final extension was accomplished at 72°C for 5 minutes. Optimal annealing temperature of 54˚C was obtained by executing the PCR optimization by running the PCR program with annealing temperatures of 50°C, 52°C, 54°C and 56°C with the same samples for comparison.

Of the cytochrome *b* gene, a 719 base pairs (bp) long partial segment, covering the end part of the 1140 bp long gene, was amplified by primers Ept_L486 and H15395_pip (Figure 2.3), since the whole gene in the museum samples failed to amplify consistently in one piece.

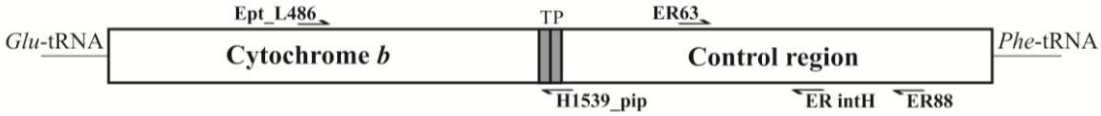


Figure 2.3 Primer locations on cytochrome *b* gene and control region (HVII). Ept_L486 and ER63 are forward and H15395_pip, ER intH, ER88 are reverse primers.

Of the control region, HVII was amplified using two sets of primers: (i) ER63 with ER88 amplified an over 500 bases long sequence and (ii) a pair with ER63 and ER intH amplified a sequence of approximately 330 bases long (Figure 2.3). The longer over 500 bp sequence (ER63/ER88 pair) failed to amplify in the museum samples, but was successfully obtained from most of the fresh tissue samples.

2.5 Electrophoresis of PCR products

The PCR products were electrophoresed in 1.5 % agarose gel and the results of the run were checked by inspecting the gel on an ultraviolet (UV) illuminated desk. Mostly 2 µl of undiluted DNA extract was used in the PCR, but when the DNA failed to amplify properly (the bands were dim or not visible on the gel), the DNA concentration in the extracts were measured using Eppendorf BioPhotometer. If the concentration was considerably high, the DNA sample was diluted to 1:10 with water, of which 2 µl were added into the reaction mix. If the DNA concentration in the extracts was very low, 4 µl or 6 µl of undiluted DNA extract was used, depending on the dimness of the band on the gel. For keeping the total volume of the mix constant, water volume was reduced in the same proportion as DNA extract was added.

Some of the samples, that were low in DNA concentration and did not amplify properly, were taken from the gel by hand using 50–1000 µl pipette tips. A few well amplified samples, visible on the UV illuminated desk, were used as a meter for finding the very dim or invisible bands. The pipette

tips were shortened by a few millimeters using disinfected scissors to obtain bigger pieces of the gel. The DNA band was collected quickly off the gel to avoid shattering of DNA caused by the UV light. The pipette tips containing the collected pieces of the gel were placed into Eppendorf tubes filled with a small quantity (50–100 µl) of water and were incubated in room temperature over night. The DNA acquired from the gel was re-amplified by PCR using the same primers and the same program to obtain a higher amount of product. The PCR results were again checked on the UV illuminated desk after electrophoresis.

2.6 Sequencing

Amplification products were prepared for sequencing in two stages: first running a sequencing reaction in PCR using the BigDye® Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems), and then cleaning the PCR products with Sephadex columns (Appendix D). For the sequencing reaction, each primer was pipetted independently in tubes in order to create separate single-stranded forward and reverse strands. The prepared sequencing reactions were submitted for sequencing runs to the Helsinki University MES-laboratory in Viikki (Biocenter 3), where they were resolved by capillary electrophoresis using the *MegaBACE 1000 DNA Analysis System* automatic sequencer. The sequencing results were received electronically from Viikki for further editing and use.

2.7 Data analyses

The forward and reverse DNA strands of both markers were cleaned and trimmed visually, and consensus sequences were created in Sequencher 4.8. The cytochrome *b* and HVII sequences were mostly analyzed separately since these mtDNA segments evolve at different rates (see section 3.3), and for cytochrome *b* each codon positions as well as transitions and transversions were considered in the analyses. In addition to the data generated in this study, previously published sequences from Artyushin *et al*. (2009) and Ruedi & Mayer (2001) were included in some of the analysis.

2.7.1 Pretreatment of the data

Some of the sequence data of this study could represent close relatives because of the colonial lifestyle of the Northern Bat females and their pups during summers (Dietz *et al*. 2009). All individuals sampled from a colony have an identical location and a few of the individuals from museum collections had also been found from a single site. The markers, cytochrome *b* and HVII, used in this study have a resolution that fails to identify closely related individuals from each other and thus they have identical sequences. To avoid bias in the estimates of diversity, the inferred close relatives were removed from the data. Given the high overall haplotype diversity, individuals with identical DNA sequences from the same collection location and date were assumed to be close relatives, and only one individual of such groups of relatives was kept in the data. The Swedish data could not be operated in this way because of incomplete information of the museum specimens; only the collection locations were received with the tissue samples. Museum collections are formed over long periods of time, because collections of bats often consist of specimens found and sent by the public (active collection by the museum staff is prohibited by conservational reasons), and thus it is possible that the individuals are not related.

In practice, initial phylogenetic trees built separately of each of the total cytochrome *b* and HVII datasets were constructed with MEGA 5 (Tamura *et al*. 2011) to help to discover the identical sequences. The assumed close relatives were excluded from all of the following analyses, except from intercalibration of mitochondrial markers and from the examination of the Finnish colonies.

2.7.2 Phylogenetic reconstruction

The genealogical history of the Northern Bat mitochondrial molecule was illustrated using the neighbor-joining (NJ) method, which produces a single tree of sequence relationships, and is based on a pairwise genetic distance matrix. It is not guaranteed that the tree represents the real genealogy, but simulations have shown that it obtains the correct topology quite efficiently. The NJ method allows the rate of molecular change to vary between branches. The tree produced is unrooted and it does not assume an evolutionary clock. The ingroup tree can be rooted by including an outgroup (Saitou & Nei 1987).

For clearer presentation of the tree topology, trees were constructed from data that only comprised the distinct haplotypes, i.e. excluded any replicates of a certain sequence type. The cytochrome *b*

sequence data of the Northern Palearctic scale and HVII sequence data of the North European scale were collapsed to haplotypes. The trees were built separately for each marker with the NJ method with MEGA 5 (Tamura *et al*. 2011) for the number of nucleotide differences, i.e. uncorrected distances.

The tree for cytochrome *b* data was built for a general view of within-species relationships of the Northern Bat, and was rooted by Serotine sequences. The HVII sequences were used to observe the intraspecific variation of the Northern Bat more closely. For HVII, the root of the tree was deduced from cytochrome *b* tree because of the lack of Serotine HVII sequences.

2.7.3 Intercalibration of mitochondrial markers

Different segments of the mitochondrial genome evolve at different rates depending on the function of the segment (Avise 2004). The protein coding cytochrome *b* gene has a relatively slow substitution rate because of functional constraints, whereas changes in the HVII segment of the control region are more frequent (Irwin *et al*. 1991; Sun *et al*. 2009). The conventionally used evolutionary rate of change of mammalian mtDNA is 0.01 substitutions per base pair per lineage per million years (μ), or 2 % sequence divergence in one million years (Myr) between pairs of lineages (Avise 2004), and this rate will be applied to the cytochrome *b* sequences, but there is no general estimate for the rate of nucleotide substitution of HVII that could be credibly applied to the Northern Bat. An estimate for the HVII rate was therefore obtained here by comparing the divergence estimates from the HVII and cytochrome *b* segments across the same data sets.

Pairwise *p*-distances corrected for multiple substitutions in a nucleotide site by a Jukes-Cantor model (explained in detail in the next section) for each pair of individuals were first calculated separately for both markers. In these data sets sequences of closely related individuals were included in the analysis. The calculations were performed with MEGA 5 (Tamura *et al*. 2011). The entries in the two distance matrices were then compared in a regression analysis. The HVII distances were set as the dependent and cytochrome *b* distances as the independent variable. A linear regression forced to pass through the origin was performed with SPSS 20 to obtain an estimate of the ratio between the substitution rates of the two gene segments. The comparison was also illustrated by a scatterplot with a regression line.

2.7.4 Estimation of divergence time

Approximate dates for events in species history can be estimated from sequence information, if the evolutionary rate of the molecular marker is known (Bromham & Penny 2003). As with the intercalibration of mitochondrial markers, a 2 %/Myr sequence divergence rate for the total bat mtDNA and cytochrome *b* was assumed here.

Corrected values of sequences were applied for time estimation. This is because the simple observed or *p*-distance does not take into account multiple mutations occurring at the same nucleotide site, and therefore gives underestimates of the true amount of change. It is defined as the proportion of nucleotide sites at which two sequences are different $p = n_d/n$, where n_d is the number of nucleotide differences between two sequences and n is the length of the sequence (Nei & Kumar 2000). The longer the time since two species have been diverged from a common ancestor, the greater the probability that back mutation at the same nucleotide site covers the earlier substitution making the site appear undifferentiated, when it actually has changed. The simplest model to account for this is the Jukes and Cantor nucleotide-substitution model, which assumes that nucleotide frequencies for each base in DNA are equal and that substitution rates are equal for each base. Under this model a corrected estimate (JC distance) is obtained by $K = -3/4 \ln (1 - 4p/3)$ (Jukes & Cantor 1969).

The Northern Bat individuals were grouped according to the distinct biogeographical lineages. The average between- and within-group *p*-distances were calculated for the Northern Palearctic cytochrome *b* data with MEGA 5 (Tamura *et al*. 2011). The obtained average between- and within-group *p*-distances were corrected manually with the equation above. The sequences of possibly closely related individuals were removed from the data, because addition of identical sequences to the data lowers the obtained JC distances (Hamilton 2009). For additional information and comparison the *p*-distance values from Table 3 of Artyushin *et al*. (2009) were also used and the values were transformed to JC distances. The additional sequence data retrieved from GenBank for this study was incomplete, because many of the sequences in GenBank from Artyushin *et al*. (2009) only comprised the beginning of the cytochrome *b* gene, and therefore were not directly comparable with the sequences of this study which were from the end part of the cytochrome *b*.

The estimated time point for the divergence of the Northern Bat lineages was then calculated using the JC distances by $t = p \times 0.5 \times 10^8$(Avise *et al*. 1988). The estimate represents the divergence time at the mtDNA sequence level, which results in earlier estimate than at the population level

since substitutions accumulate before the divergence at the latter level (gene tree versus species tree; Hamilton 2009).

2.7.5 Genetic diversity and phylogeographic analyses

The levels and patterns of variation of the Northern Bat mtDNA in Northern Europe were inspected from the HVII data at two geographical scales: North Europe as a whole (all Fennoscandia and Latvia), and Finland alone. Two approaches were used: building haplotype networks and estimation of molecular diversity indices and components within and between geographically defined sets of specimens.

Haplotype networks: A haplotype network is a visual illustration of the (mutational) relationships between haplotypes. Unlike simple phylogenetic trees, illustrations such as haplotype networks can present a spider web-like mesh, which represents unclear relationships caused by back mutations for instance. It represents another way to inspect genealogical information within the mitochondrial sequence data. In a statistical parsimony analysis applied here, the program first calculates frequencies of the haplotypes in the sample and then builds the network based on a matrix of all pairwise comparisons of the haplotypes (Clement *et al*. 2000). By combining the network with geographical locations of the individuals, the geographical distribution of variation can be illustrated and used for inferring the biogeographic population history.

Haplotype networks were built separately for the Northern Bat HVII data at the North European and Finland scale. The Finnish data were arbitrarily divided into four sections based on a south–north axis with approximately equal number of individuals in each section. The analyses were executed with TCS 1.21 (Clement *et al*. 2000) using the default 95 % statistical parsimony connection limit. Due to the presence of two strongly diverged lineages in the data, TCS produced two separate haplotype networks for each data set, which were not connected to each other. The networks were graphically finished in CorelDraw X5. The two separate networks were subsequently connected by a dashed line showing the smallest number of differences between the lineages.

*Components of molecular diversity***:** Estimates of nucleotide diversity (π), haplotype diversity (h), AMOVA (Analysis of Molecular Variance) and Φ_{ST} values were calculated to measure genetic variation and population divergence among the Northern Bat.

The nucleotide and haplotype diversities measure molecular variation at the nucleotide and haplotype (or allele) levels, respectively. Nucleotide diversity (π) is an expression of the degree of polymorphism, i.e. the average number of nucleotide differences per site, between sequences of two random individuals from the same populations. It is estimated by $\pi = \sum_{i<j} \pi_{ij} / n_c$, where π_{ij} is the proportion of different nucleotides in sequences of individuals *i* and *j*, and n_c is the total number of sequence comparisons (Nei 1987). Thus, the diversity for population sample of *n* individuals is obtained as an average of the pairwise values. Nucleotide diversity can be used to compare the extents of polymorphism of different genes and even across species, since it should not reflect the length of the DNA sequence or sample size (Hamilton 2009). However, π is affected by the population history since it measures the degree of genetic divergence between the alleles in a sample (Avise 2004).

Haplotype diversity "condenses information on the numbers and frequencies of different alleles at a locus" (Avise 2000). It is defined as the probability that two haplotypes randomly sampled from a population are not identical: $h = 1 - \sum x_i^2$ where x_i is the frequency of the ith haplotype in a population (Nei 1987). The haplotype diversity does not consider the magnitude of genetic divergence between sampled alleles, only the number and frequencies of different observable variants. The loci used and species examined have an effect, but within-population nucleotide diversities are often in the range of 0.001–0.020 and haplotype diversities may exceed 0.5 for rapidly evolving mammalian mtDNA (Avise 2004).

Nucleotide and haplotype diversities were calculated for HVII data at the North European scale for both lineages (European and Siberian), for each country individually (Finland, Sweden, Norway and Latvia) and then separately at the Finland scale. Since there are no obvious geographical boundaries preventing bat dispersal within Finland, it was subdivided arbitrarily to geographical latitudinal groups in order to examine if there are any population structure present. The genetic diversity calculations were executed with DnaSP 5.10 (Librado & Rozas 2009).

The hierarchical division of diversity into intra- and inter-population components was also presented in an analysis of molecular variance (AMOVA). AMOVA can be performed in terms of Wright's *F*-statistics for the haplotype diversity data or their analogues, *Φ*-statistics, for nucleotide diversity. The latter was applied here; in analogy with F_{ST}, the inter-population component Φ_{ST} estimates the proportion of nucleotide diversity among subpopulations, relative to the total variance (π_T). The significance of the *Φ*-statistics, i.e. population differentiation, was tested by a permutation approach (Excoffier *et al*. 1992).

In principle, AMOVA can be performed simultaneously for multiple hierarchical or nested levels of diversity or population structure. Here it was applied simply to single level geographical subdivisions, separately for the North European data set (subdivided to countries) and to the Finland-only data set (subdivided to arbitrary latitudinal sectors, as above; Table 2.2). This is justifiable in a situation where the within-Finland variation turned out to be non-significant and negligible (and thus should not affect the variance components in the analysis of a broader scale).

Table 2.2 Data subdivisions of the Northern Bat individuals (*I*) according to the scales of North Europe and Finland.

Scale	**Subdivision**	***I***
Northern Europe (*N* = 109)	Finland (including Åland)	65
	Sweden (including Gotland)	24
	Norway	7
	Latvia	13
Finland (*N* = 65)	North	15
	Middle	16
	South	16
	SW coast, Åland	18

Close relatives and the Siberian lineage, because of its evidently different origin than that of the European lineage, were excluded from the analysis. The Siberian lineage was not examined separately because of the small sample size. AMOVA was executed with Arlequin 3.1 (Excoffier *et al*. 2005; testing with 1000 permutations). Pairwise population comparisons, in terms of Φ_{ST}, were also calculated and tested from the same data sets (Table 2.2) with Arlequin (500 permutations and significance level 0.05).

The composition of individuals in the sampled five Finnish Northern Bat colonies (Figure 2.4) was found interesting, and an NJ tree was built of the HVII data with MEGA to illustrate the division of genetically different individuals in each colony. Closely related individuals were included in the data (listed in Appendix A).

Location	*N*
Savuniemi, Joroinen	9
Kerisalonsaari, Joroinen	6
Hamina	11
Siuntio	6
Loviisa	6
Total	38

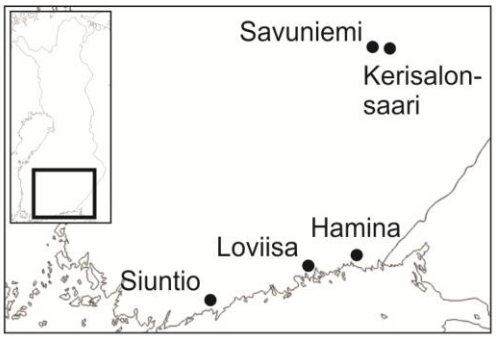


Figure 2.4 Number of individuals (*N*) in each sampled Finnish Northern Bat colony and a map of colony locations in southern Finland.

2.7.6 Distribution of pairwise differences and dating of demographic events

Distribution of observed pairwise nucleotide site differences (or individual *p*-distances), i.e. mismatch distributions, may bear signatures of past demographic events. Particularly an instance of fast population growth following a bottleneck creates a wave to the histogram of pairwise differences. The crest of the wave represents the time since expansion in mutation units $\tau = 2ut$, in which t is time in years and u is the substitution (mutation) rate of the entire segment of mtDNA under study, $u = n_c\mu$, where n_c is the sequence length and μ is the mutation rate per site. The resulting values can be used to estimate when the observed demographic events occurred (Rogers & Harpending1992).

Mismatch distribution histograms were created separately for the North European Northern Bat cytochrome *b* and HVII data. The observed distributions were fitted with expected distributions generated by simulation of a population growth-decline scenario, using DnaSP 5.10 (Librado & Rozas 2009). The graphs obtained with DnaSP were finished into histograms with Microsoft Office Excel 2007.

The dates of the demographic events were estimated from the τ represented by the mismatch peaks ($t = \tau/2u$). The per site substitution rates 2μ = 2 %/Myr of cytochrome *b* and that estimated for HVII in section 3.4 were then used to obtain *u* and the estimated dates of population expansion in years.

2.7.7 Parametric and Spearman's rank correlations for genetic diversity and latitude

Decreasing northward genetic diversity is often detected in populations which have expanded to north from their southern refugia, because usually only a few founder individuals dispersed from the source population in a series of expansions (Hewitt 1999).

To look for such a pattern, the North European Northern Bat individuals were first arranged by the latitudes obtained for the location of each individual with Google maps (http://itouchmap.com/latlong.html; data presented in Appendix A), and then divided into five latitudinal groups of similar size: The first (S1) and second (S2) groups consist of the Latvian and southernmost Fennoscandian individuals, the third (M3) middle latitude Fennoscandian individuals and the fourth (N4) and fifth (N5) the northernmost Fennoscandian individuals. The group division is shown later on a map in the results in section 3.8 (Figure 3.8). The groups were first formed with DnaSP 5.10 setup, and then nucleotide and haplotype diversities were calculated for each group. The presence of a consistent latitudinal trend was assessed with Microsoft Office Excel 2007 using parametric and Spearman rank correlation coefficients between the group-wise diversity estimates and the mean latitude of the samples in each group.

3 RESULTS

3.1 Sequence analyses

Sequencing was successful on a total of 146 of the 151 sampled northern Palearctic Northern Bat individuals, and thus only from 5 individual museum specimens sequences were not obtained. All fresh tissue samples were successfully sequenced. The HVII sequence was obtained from 145 individuals and cytochrome *b* from 92 (listed in Appendix A).

The length of the cytochrome *b* fragment was trimmed to match the length of additional sequences of the Northern Bat and Serotine, which were obtained from GenBank (Appendix B), and thus set at 603 bp covering approximately 55 % of the total gene length from the end part of the gene. The HVII sequences in turn were trimmed to a final length of 327 bases between the primers ER63 and intH. The sequence further downstream between primers intH and ER88 (see Figure 2.3) comprised a repeated sequence motif, in which the first 40 bp were CRYRTA (R = G/A, Y = C/T), and the rest over 100 bp consisted of CGCATA alternating with CGTGTA (data not shown). These repeats occurred in varying number (approximately 15–24), and thus could not be used in standard analyses based on unequivocal alignment. This repeat motif has been reported also in *Nyctalus noctula* (Petit *et al*. 1999) and in Myotis myotis, in which, as in the Northern Bat, two different forms exist: CGCATA (Castella *et al*. 2001) and CGTATA (Petri *et al.* 1996).

The trimmed sequences were aligned in BioEdit version 7.0.5.3 (Hall 1999). There were no insertions or deletions in the aligned sequences (Appendix E). From the cytochrome *b* fragment 14 different haplotypes were defined that were characterized by 24 variable sites of which most were transitions. Position 129 is the only exception where the most common cytosine was replaced by a guanine in one bat (Appendix E, Table 1). The HVII sequences collapsed into 79 haplotypes, which are characterized by 48 transitions and by a single transversion in position 289 in one bat (Appendix E, Table 2). At this position the most common bases guanine and adenine are replaced by thymine.

The sequence data were distributed to the different embedded geographical (and systematical) scales in the following manner: The largest scale, the northern Palearctic, consists of 111 cytochrome *b* sequences, of which 92 are from this study, and 6 Northern Bat sequences as well as 13 Serotine sequences (outgroup species) are from GenBank. The HVII sequences were not used for this scale. The mid-scale, North Europe, comprises only Northern Bat sequences: 145 sequences of HVII and the same data from previous scale for cytochrome *b* (98 sequences), but without the

Serotine sequences. The most local scale, Finland, contains 101 HVII Northern Bat sequences (including one from Vyborg, Russia). These amounts of sequences were retrieved from the complete data before any exclusion of closely related individuals.

3.2 Mitochondrial lineage history of the Northern Bat and outgroup species

The mitochondrial lineage history of the northern Palearctic Northern Bat (*Eptesicus nilssonii*) was analyzed using the cytochrome *b* sequence data of 92 northern bats of this study as well as 6 Northern Bats and 13 Serotines retrieved from GenBank. The Northern Bat population comprises two distinct lineages. In the total analyzed cytochrome *b* data set, a total of 14 distinct Northern Bat haplotypes were observed; 11 in the first lineage, and 3 in the second one (Figure 3.1). Haplotypes of the first lineage were present only in Europe (including Fennoscandia), and haplotypes of the second lineage both in Siberia (data from Artyushin *et al*. 2009) and Finland. Artyushin *et al*. (2009) already detected the same lineages and named them respectively as the European lineage and the Siberian lineage. In more detail, the Siberian lineage clade formed mixed groups of individuals from Finland and Russia (Figure 3.1).

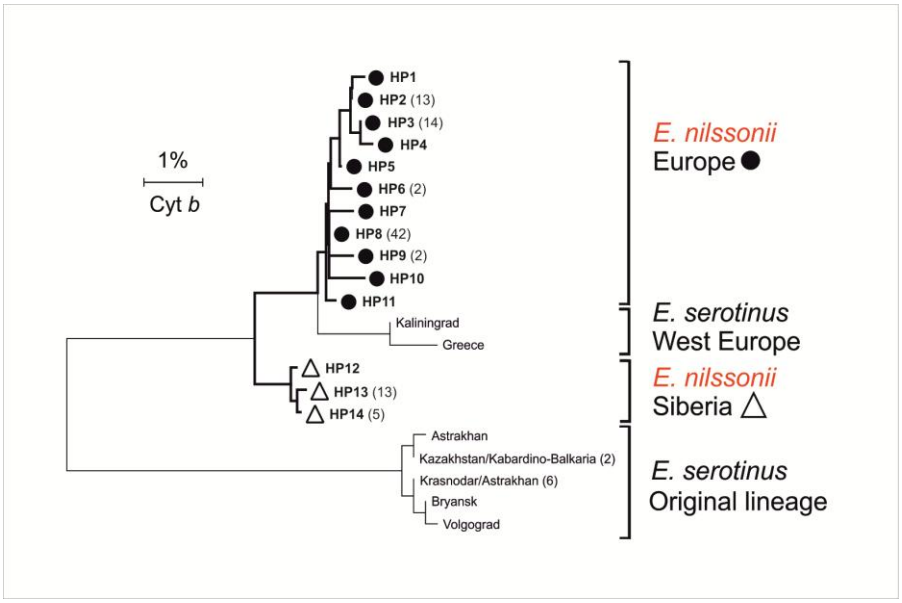


Figure 3.1 Neighbor-joining haplotype tree of the Northern Bat (*Eptesicus nilssonii*) and Serotine (*E. serotinus*) cytochrome *b* (cyt *b*) sequence data comprising of only distinct haplotypes. The numbers of individuals are shown in brackets whenever there are more than one individual in the haplotype; detailed information of the individuals belonging to each haplotype can be found in Appendix A and B.

The diversity of the Serotine (*E. serotinus*), which was used as an outgroup, also comprised two lineages. The majority of the Serotine haplotypes show considerable difference from all Northern

Bat haplotypes, but the two westernmost haplotypes from Kaliningrad and Greece were more closely related to all of *E. nilssonii* and especially to the European lineage of the Northern Bat than to the five haplotypes of the more easterly *E. serotinus* (Figure 3.1). These Serotine groups were named respectively as the original mitochondrial lineage and the West European mitochondrial lineage by Artyushin *et al.* (2009).

The same two main Northern Bat lineages were found in the North European HVII sequence data, which both showed generally more extensive intraspecific diversity than with cytochrome *b*. Of a total of 65 distinct Northern Bat haplotypes, observed among 119 HVII sequences, 56 haplotypes were of the European lineage and 9 haplotypes of the Siberian lineage (Figure 3.2).

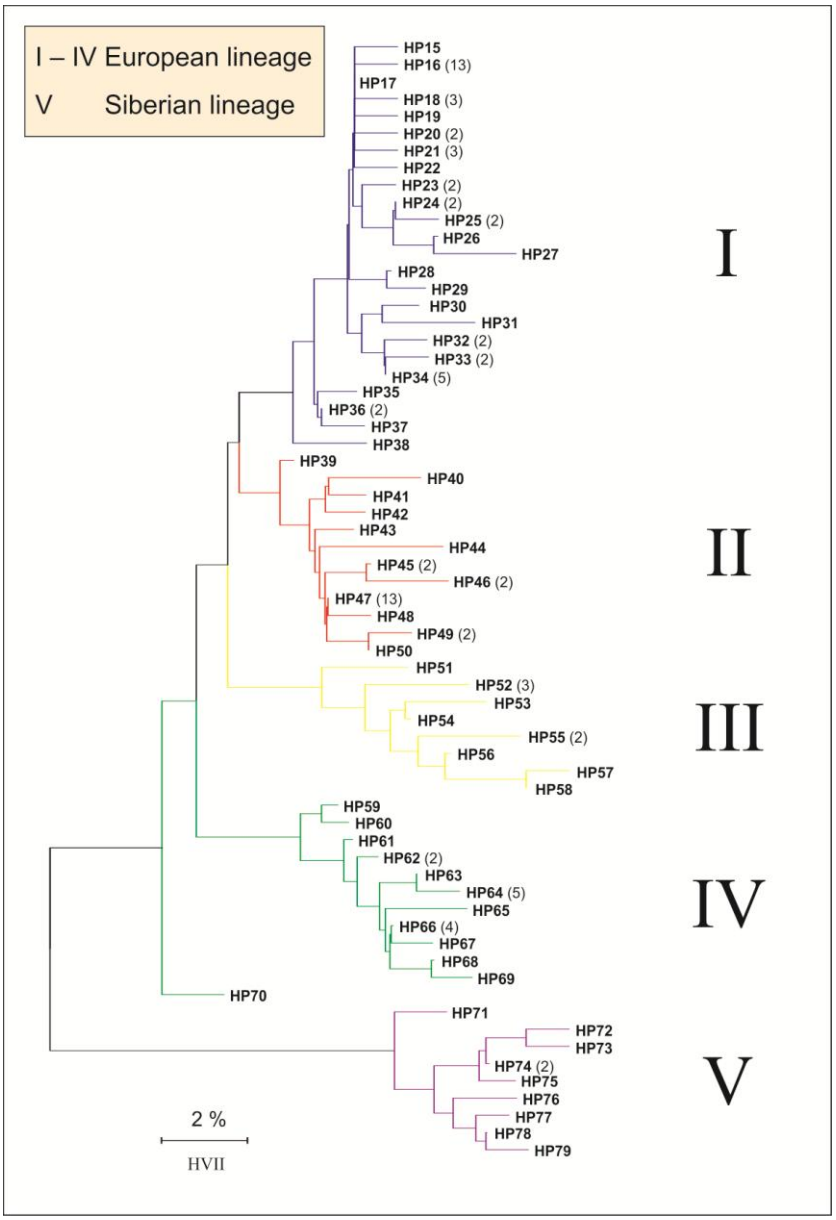


Figure 3.2 Neighbor-joining haplotype tree of Northern Bat HVII sequence data showing clusters I–V. The numbers of individuals in each haplotype are in brackets whenever a haplotype consist more than one individual. More information of the individuals in each haplotype can be found in Appendix A.

The HVII genealogy showed further structuring: the branches of the NJ tree could visually be divided into (at least) five clusters, I–V. Of these, I–IV represent the European lineage and V the Siberian lineage.

3.3 Evolutionary rates of mitochondrial markers

The regression of the mitochondrial markers was based on data of this study, which comprised 84 sequences of the same individuals for each marker. From the regression of corrected pairwise differences of cytochrome *b* on those of HVII, the average differentiation in HVII sequences was found to be 3.5 times higher than in cytochrome *b* (unstandardized regression coefficient $b = 3.46$, $P < 0.001$). The two variables were significantly correlated ($R = 0.902$, $P < 0.001$), and cytochrome *b* explained 81.4 % of HVII (Figure 3.3). If the cytochrome *b* evolves at 2 %/Myr, then the rate of divergence for the HVII was estimated to be 7 % per million years.

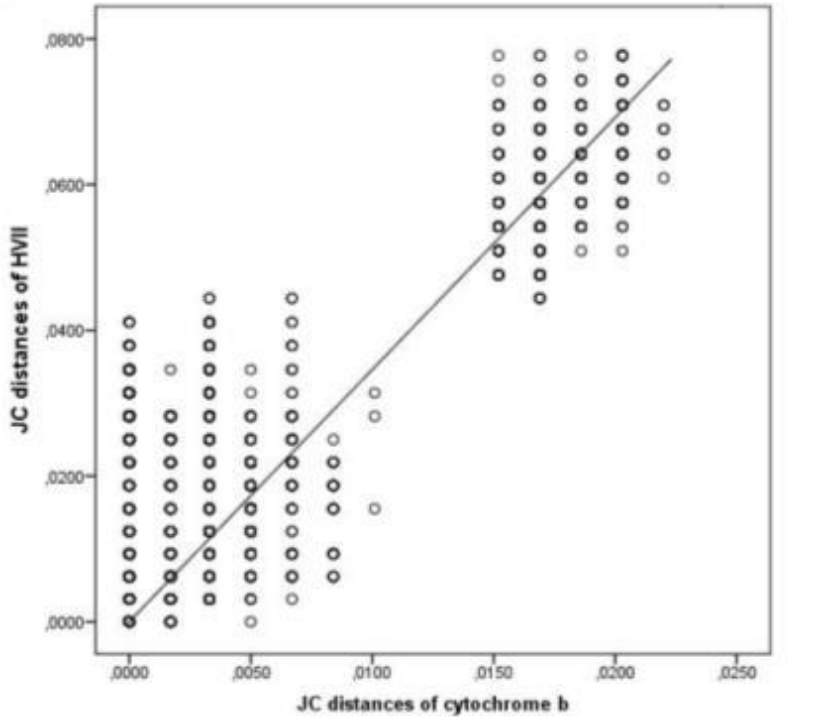


Figure 3.3 Scatter plot for pairwise JC distances of cytochrome *b* and HVII (linear $R^2 = 0.704$).

3.4 Divergence time of the Northern Bat

The divergence time between lineages was estimated from the Northern Palearctic cytochrome *b* data on the basis of the 2 %/Myr mammalian rate, which was applied to JC-corrected average inter-lineage distances. The JC-corrected distances were calculated from the cytochrome *b* data consisting of 65 Northern Bat sequences from this study as well as 6 Northern Bat sequences from GenBank. In addition to the distances from the end part of the cytochrome *b* (603 bp) alignment of this study, distances based on the largely non-overlapping beginning part segments and the complete gene varying from 726 bp to 1140 bp alignment from Artyushin *et al*. (2009) are shown in Table 3.1. Distances from the different data sets are quite similar. Within the lineages, distances of the European lineage are higher than those of the Siberian lineage, and the highest distances were found between the lineages.

Table 3.1 Estimates of JC-corrected average inter-lineage (below diagonal) and within-lineage (diagonal) cytochrome *b* distances of the Northern Bat (European and Siberian lineage) from this study (including aligned sequences from Artyushin *et al*. 2009 and Ruedi & Mayer 2001 used in this study (Appendix B); bold numbers) and from the complete data of Artyushin *et al*. (2009; lower numbers).

Lineage	European	Siberian
European	**0.003**	
	0.006	
Siberian	**0.018**	**0.0009**
	0.022	0.0004

Divergence of Northern Bat lineages: The distance between the Siberian and European Northern Bat were 1.7–2.2 %, corresponding to 0.85–1.1 Mya with the standard calibration.

3.5 Geographical differentiation among North European populations

The subdivision of the HVII genealogy into clusters is also illustrated by the haplotype network consisting of 119 North European Northern Bat sequences (Figure 3.4). Clusters I and II form a star-phylogeny, i.e. a common haplotype, and a set of neighboring haplotypes that are separated from the most common one by a single mutation (Harpending *et al*. 1998). Clusters III, IV and V are not as clearly star-shaped. Each cluster in the European lineage is distinctly separated from the others by several mutations (steps).

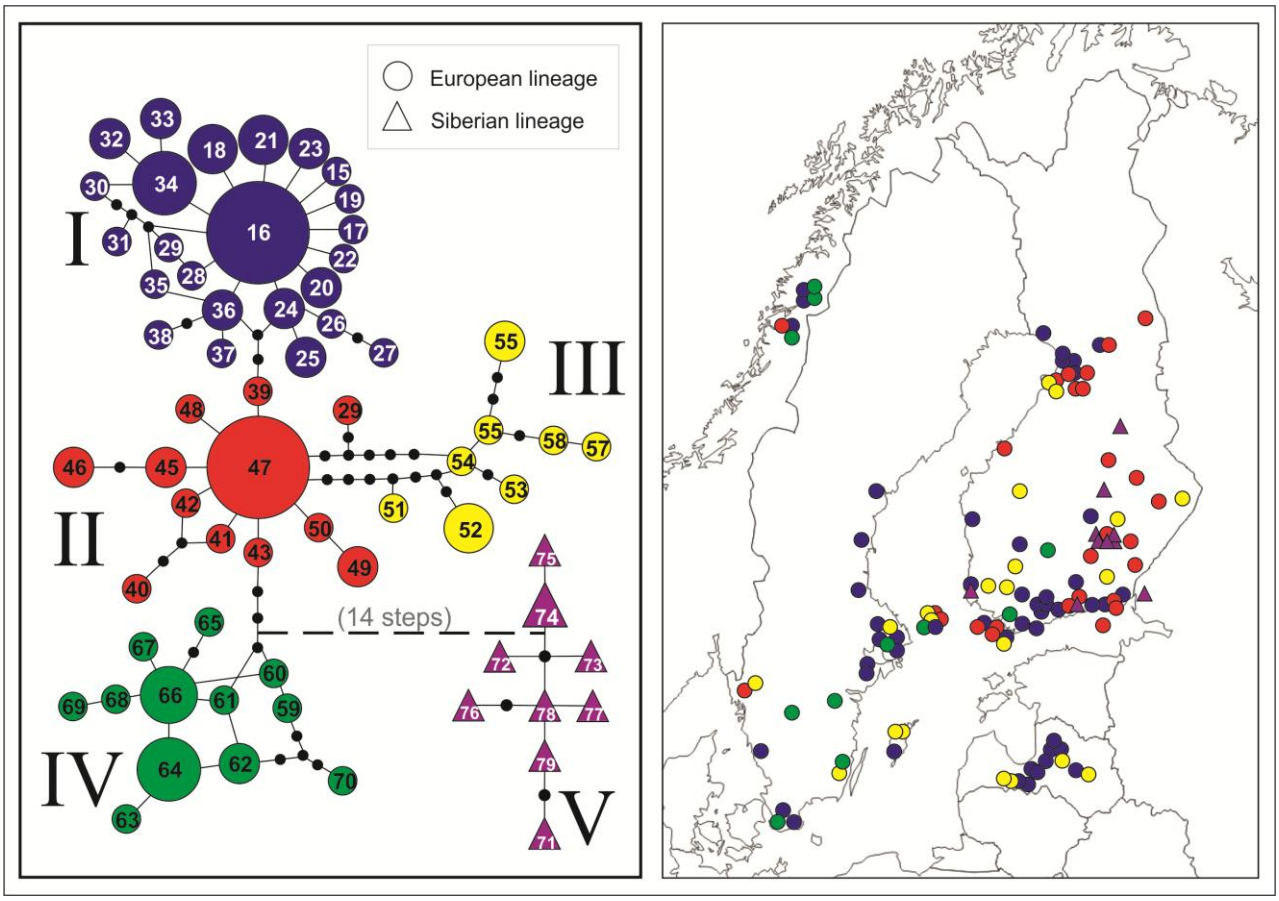


Figure 3.4 Haplotype network and corresponding sample sites of Northern Bat individuals on the map. The dashed line between clusters IV and V illustrates the arbitrarily located connection between the European and Siberian lineages. The sizes of dots and triangles represent the relative number of individuals representing each haplotype in the data. The lines connecting the dots represent single mutations, and the small black dots are hypothetical intermediate haplotypes, which are not present in the data. The smallest colored dots and triangles represent a single individual, the largest dot 13 individuals, and the largest triangle two individuals. The numbers within dots and triangles correspond to the haplotype numbers in Figure 3.2.

Plotting the clusters I–V on the sample map suggests an uneven geographical distribution of these clusters within northern Europe (Figure 3.4). Cluster I was found from all countries sampled, but clusters II and IV were absent in Latvia and Gotland, and cluster III was absent in Norway. The individuals of Siberian lineage, cluster V, are present only in Finland (including a sample from Vyborg) and are clustered but not restricted to south-eastern Finland. Individuals of the Siberian and European lineages were found at the same locality and even in the same bat colony (Joroinen: Savuniemi).

The overall nucleotide diversity of the European lineage is higher than that of the Siberian lineage, which is represented by only a single cluster (Table 3.2). The haplotype diversity values were similar in both lineages.

Table 3.2 Overall nucleotide diversity (π) and haplotype diversity (h) estimates for the Northern Bat European lineage and Siberian lineage HVII data, and the number of sequences analyzed in each cluster (N).

Lineage	π	h	N
European	0.019	0.966	109
Siberian	0.009	0.978	10

AMOVA suggests statistically significant differentiation among countries ($\Phi_{ST} = 0.07$) in the European lineage ($N = 109$).

In pairwise Φ_{ST}'s, differences of the values in comparisons between Finland, Norway and Latvia were relatively high and significant (Table 3.3), whereas the comparisons with Sweden were low and non-significant. The Siberian lineage was removed because of the high sequence divergence between the lineages.

Table 3.3 Pairwise Φ_{ST}'s of the HVII data that was grouped by country for the Northern Bat European lineage. Statistically significant values are marked with an asterisk.

Country	Finland ($N = 65$)	Norway ($N = 7$)	Latvia ($N = 13$)
Norway	0.12*		
Latvia	0.12*	0.18*	
Sweden ($N = 24$)	0.03	0.04	0.02

*P-value <0.05

In further examination of within countries diversity, the nucleotide and haplotype diversities were the highest for the population in Sweden (Table 3.4). Nucleotide diversities were of middle level in Finland and Norway, and the lowest in Latvia. The haplotype diversity in Finland was nearly equal with that of Sweden, middle scale in Latvia and the lowest in Norway. The Siberian lineage was again excluded from the data.

Table 3.4 Within country nucleotide and haplotype diversities for the European lineage data, and the number of individuals (I) in each country.

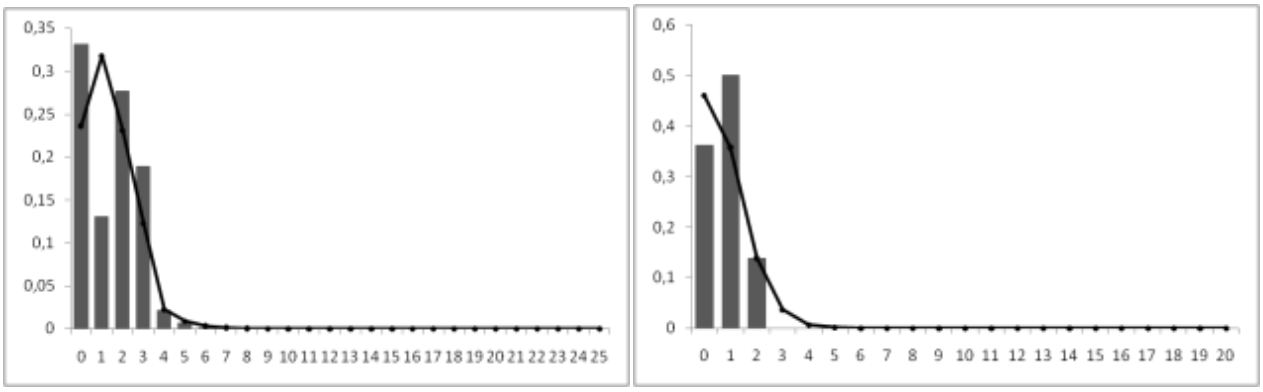
Country	π	h	I
Finland	0.01766	0.956	65
Norway	0.01718	0.857	7
Latvia	0.01451	0.910	13
Sweden	0.02183	0.957	24

3.6 Signatures of past demographic events

The demographic event estimation was based on HVII (*N* = 119) and cytochrome *b* data of this study (*N* = 65, of which 59 in the European and 6 in the Siberian lineage) and from GenBank (*N* = 6, of which 2 in the European and 4 in the Siberian lineage). The mismatch distributions, constructed separately for the European and Siberian lineages on the North European scale, showed similar patterns for both mitochondrial lineages whether they were based on the cytochrome *b* or HVII data. The European lineage has two peaks, and the Siberian lineage has a single peak (Figure 3.5).

Cyt *b* European lineage (*N* = 61) Cyt *b* Siberian lineage (*N* = 10)

τ = 0.969 τ = 0.778



HVII European lineage (*N* = 109) HVII Siberian lineage (*N* = 10)

τ = 3.892 τ = 2.844

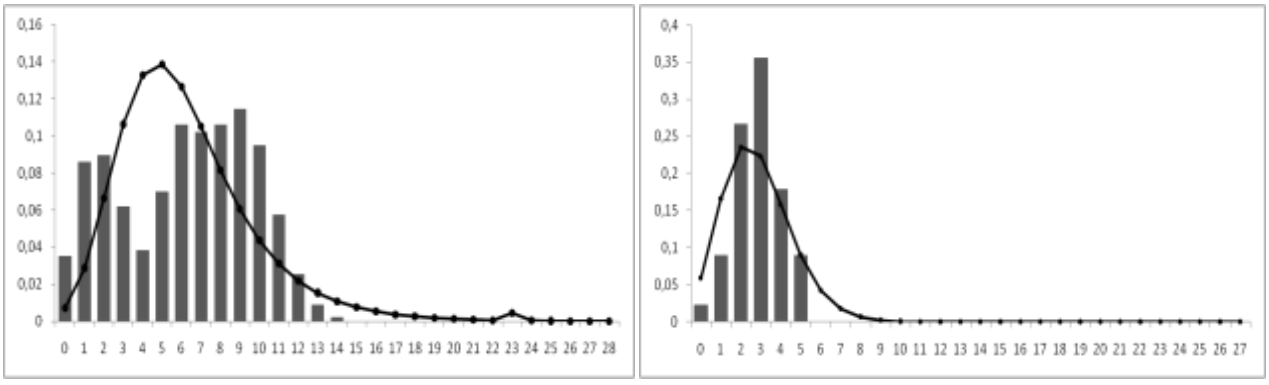


Figure 3.5 Mismatch distributions of the Northern Bat mitochondrial lineages for both markers. The bars represent observed values and the line represents expected values of the sudden expansion model, estimated by DnaSP.

Consequently, the distributions of the European lineage fitted poorly to the simple expansion model, which predicted a single peak for a sudden expansion situation, as well as with cytochrome *b* the Siberian lineage, in which the model predicted a reduction in population size (L-shaped

curve). For the two-peak distributions of both markers and the one-peak distribution of cytochrome *b* distribution, the time of expansion was simply estimated from the observed peak; the rate estimates from section 3.3 were used here. The estimates are shown in Table 3.7.

Table 3.7 Estimates of the time of population expansions in the Northern Bat European lineage (Eur 1 and 2) and Siberian lineage (Sib) based on the approximate value of τ.

				τ (observed crest)			***t* = τ/2*u* [Myr]**		
marker	**rate 2μ (Myr)**	**length n_c (bp)**	**rate 2*u* = 2μ*n_c (Myr)**	Eur 1	Eur 2	Sib	Eur1	Eur2	Sib
cyt *b*	0.02	603	12.06	0	2	1	0.00	0.17	0.08
HVII	0.07	327	22.89	2	8	2.8	0.09	0.35	0.12

The time estimates suggest that the population expansions occurred in the Holocene to Middle Pleistocene (0.0–0.781 Mya; Cohen *et al*. 2013).

3.7 Genetic variation in the Finnish Northern Bat population

Individuals from the four arbitrarily defined latitudinal subdivisions within Finland (color-coded on the map, Figure 3.6) seem to be similarly divided among a few clusters of the haplotype network. However, the haplotypes of the Siberian lineage are found mostly in central Finland (green and orange) and less in northern Finland (red), but are absent in the southernmost Finland and Åland (grey).

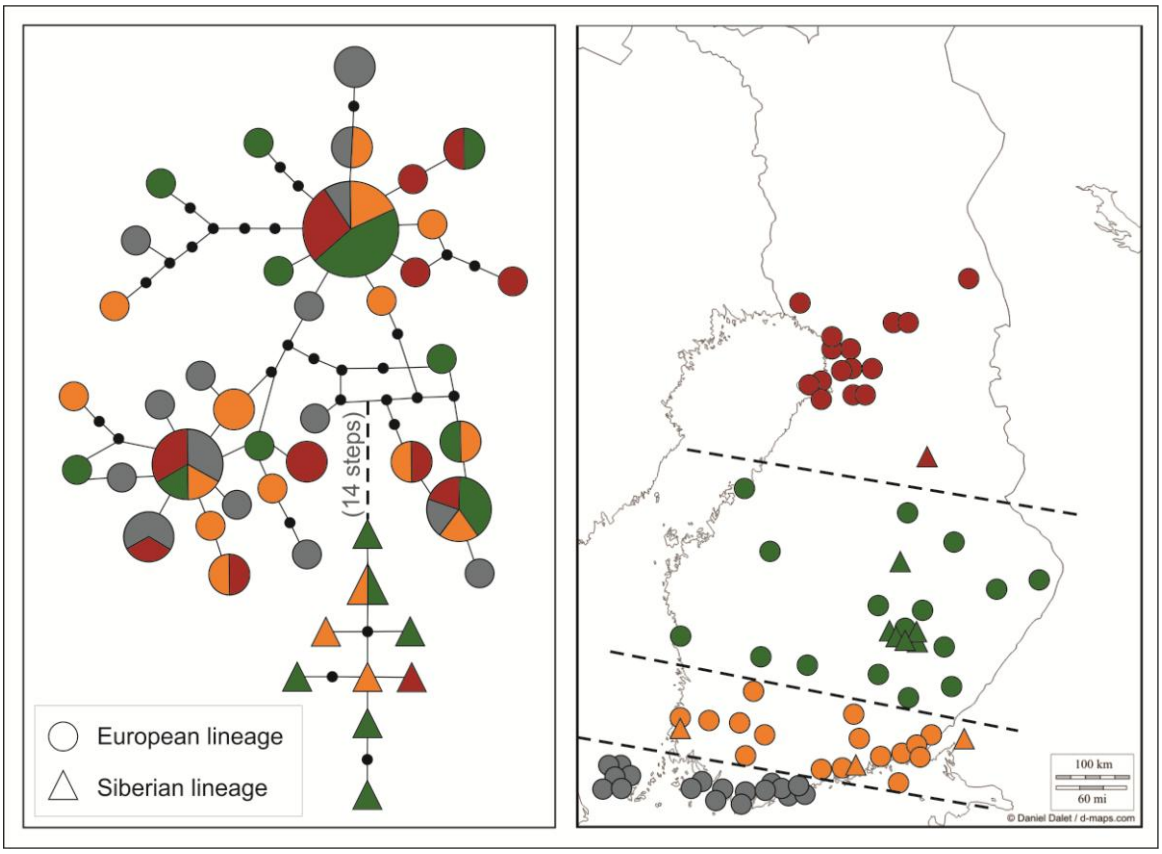


Figure 3.6 HVII haplotype network and sample sites of Finnish Northern Bats. European lineage comprises 36 distinct haplotypes (from *N* = 65 individuals) and Siberian lineage 9 haplotypes (from *N* = 10). The orientation of the dashed lines is arbitrary, but the number of individuals in each section is nearly equal.

Respectively, nucleotide and haplotype diversity estimates for the geographical groups were high and similar between the groups (Table 3.5) with an average of 0.018 and 0.956 of the total data (European lineage data only). These values are also similar to that of the European lineage at the north European scale (π = 0.019 and h = 0.966). The Siberian lineage was not included for reasons explained above, and diversities within that lineage were presented in the previous section.

Table 3.5 Nucleotide diversity (π) and haplotype diversity (h) estimates and group sizes (l) for geographical subdivisions of the Finnish Northern Bat European lineage data (N = 65).

Group	π	h	l
North (red)	0.016	0.952	15
Middle (green)	0.018	0.942	16
South (orange)	0.019	0.975	16
SW coast and Åland (grey)	0.018	0.980	18

AMOVA on nucleotide diversity for the European lineage revealed no appreciable differentiation among the subdivisions in Finland (Φ_{ST} = -0.002). Also none of the pairwise Φ_{ST} estimates were significant and only a single value was positive, i.e. SW coast vs. Middle Finland (Table 3.6). This

analysis again excluded the Siberian lineage, which distribution however appeared uneven within the country (Figure 3.6).

Table 3.6 Pairwise Φ_{ST}'s for geographical subdivisions of the Finnish Northern Bat HVII data (European lineage only, *N* = 65). None of the values were statistically significant.

Area	SW coast, Åland (grey)	South Finland (orange)	Middle Finland (green)
South Finland (orange)	-0.0235		
Middle Finland (green)	0.0480	-0.0054	
North Finland (red)	-0.0059	-0.0338	-0.0047

A part of the Finnish data was sampled from five different colonies, of which the colonies of Joroinen turned out to be of special interest. The colony of Savuniemi consisted of individuals belonging to both the European and Siberian lineage, and the colony of Kerisalonsaari consisted of only Siberian lineage individuals (Figure 3.7). All the other colonies comprised merely European lineage individuals.

Colony	European lineage	Siberian lineage	*I*
Savuniemi, Joroinen (A)	3	6	9
Kerisalonsaari, Joroinen (B)		6	6
Hamina (C)	11		11
Siuntio (D)	6		6
Loviisa (E)	6		6
Total	26	12	38

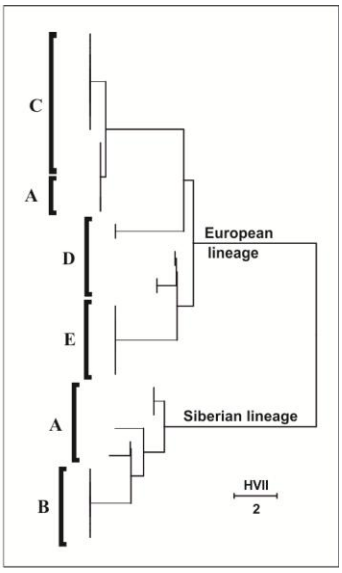


Figure 3.7 The number of individuals in colonies according to the lineage they belong to, the total number of individuals in each colony (*I*), and a haplotype tree of the HVII data, in which close relatives were included.

3.8 Correlation between genetic diversity and latitude

The correlation was based on the HVII data of 109 sequences of this study; the Siberian lineage was excluded from the analysis. The nucleotide diversity values of the European lineage Northern Bats show an apparent south to north reduction (Figure 3.8) for the North European transect, but the

trend for the haplotype diversity is less clear. The correlation between nucleotide diversity and group average latitude was r = -0.91 (d.f. = 3, P< 0.05, two tailed; Spearman ρ = -1), and for haplotype diversity the values were also negative but not significant (r = -0.61, d.f. = 3, P> 0.05, ρ = -0.60).

Group	π	h	l	lat ($\bar{x}$)
N5	0.0177	0.944	22	65.5
N4	0.0180	0.935	22	62.0
M3	0.0185	0.974	22	60.4
S2	0.0186	0.952	22	59.4
S1	0.0202	0.967	21	56.8

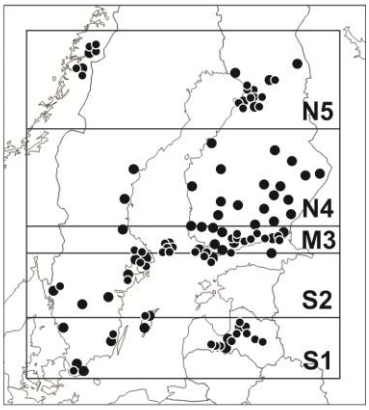


Figure 3.8 Nucleotide diversity (π) and haplotype diversity (h) values, the number of individuals (l) and average latitude (lat ($\bar{x}$)) in five latitudinal groups of the European lineage only Northern Bat HVII data (N = 109).

4 DISCUSSION

4.1 Mitochondrial DNA diversity of the Northern Bat in the northern Palearctic and Finland

Also in this study, the northern Palearctic Northern Bat (*Eptesicus nilssonii*) population, examined in three different geographical scales, was confirmed to contain two separate and relatively deeply diverged mitochondrial lineages, the European and Siberian lineages. The European lineage shares similar sequences with the Serotine (*E. serotinus*) as a result of ancient introgression, and has considerable mtDNA diversity, which was observed as four clusters that were unevenly distributed geographically in the Northern Europe. The Siberian lineage formed the fifth cluster and was located only in Finland (and Vyborg, Russia) in the data of this study. The clusters of the European lineage were observed to have an even geographical distribution within the Finnish population, whereas the occurrence of the Siberian lineage was more localized.

At the northern Palearctic scale the European and Siberian lineages of the Northern Bat were mostly distributed to the different but partially overlapping geographical locations by which the lineages were named. Similar geographical patterns of intraspecific division in mitochondrial lineages of the

northern Palearctic bats have been discovered at least in Steppe Whiskered Bat (*Myotis aurascens*), in which the lineages were confined in Europe and Asia, and among Common Pipistrelle (*Pipistrellus pipistrellus*), in which the lineages were found in Central Europe and Caucasus, by using the mitochondrial sequences of the cytochrome oxidase I (COI; Kruskop *et al*. 2012). From the cytochrome *b* sequence data of the Northern Bat, the European and Siberian lineages were estimated to diverge in the Early to Middle Pleistocene, which occurred 2.588–0.126 million years ago (Cohen *et al*. 2013). With other mammalian species, most of the intraspecific mitochondrial lineages studied under the 2 %/Myr molecular clock were also diverged during the Pleistocene epoch (Avise *et al*.1998), even though the intraspecific lineages in two of the three bat species included in the review were later discovered to represent cryptic species (Wilmer *et al*. 1994; Barrat *et al*. 1997).

Although the division between the Northern Bat lineages is very distinct, the amount of difference (~ 2 %) between the lineages most likely reflects conspecific populations that are capable of interbreeding (Bradley & Baker 2001). Low levels (under 3 %) of divergence have been detected as interpopulation variation also among many other bat species (Hulva *et al*. 2004; Ibáñez *et al*. 2006; Bogdanowicz *et al*. 2009). Furthermore, Kruskop *et al*. (2012) did not find any detectable different morphological characters between the Northern Bat lineages.

At the North European scale, the Northern Bat HVII sequence data revealed that the European and Siberian lineages have distinct genetic characters. The European lineage comprised considerably high nucleotide and haplotype diversities. These high diversities are a signature that the mitochondria of the samples are from historically different populations (Avise 2000), and thus reflect the presence of the four clusters discovered within the European lineage. The distinction of the European lineage clusters by multiple mutations and the nucleotide and haplotype diversities already discussed also imply that these clusters represent historically sundered populations. In other European bat species (e.g. Ruedi & Castella 2003; Berthier *et al*. 2006; Bogdanowicz *et al*. 2009; Rebelo *et al*. 2012), which have clusters that were formed in separate refugia, the haplotype networks look similar to those that I obtained from my data. The Northern Bat cluster III was different from the others by having multiple steps between it and the cluster II, and between most of the different haplotypes within the cluster (Figure 3.4), thus presumably indicating that this cluster has a mixture of haplotypes formed in different ancient refugia.

On the contrary, the HVII sequences of the Siberian lineage formed a single cluster and had a nucleotide diversity value much lower than that in the European lineage. Combining this with the information of high haplotype diversity suggests that the Siberian lineage experienced a rapid

population growth from a small ancestral population size (Avise 2000), which however is not confirmed by the haplotype network, because star-phylogeny was not detected in the cluster V, which is likely due to a small sample size. Nevertheless, a single cluster origin is highly likely since, according to the haplotype NJ tree of the cytochrome *b* data, individuals of the Siberian lineage from East Russia, do not differ considerably by their mtDNA from the Finnish and west Russian Northern Bats although the distance between them is up to 5000 kilometers. The contrast between the European and Siberian lineage was also visible in the within-lineage JC distances, of which the European lineage distances were somewhat higher than that of the Siberian lineage.

Continuing at the North European scale, the mismatch distributions of both markers formed peaks in the histograms, two separate in the European lineage and one in the Siberian lineage. These waves might indicate that there have been events of rapid population growth, fast recoveries from bottlenecks. The star-phylogeny visible in the clusters of the European lineage haplotype network (Figure 3.4) is also an indication of population expansion. Steep sides of each wave signify that the original populations were small (Rogers & Harpending 1992). These small populations probably formed when, respectively, the European lineage and Siberian lineage populations were forced to isolate into multiple and single refugial subpopulations during glacials.

The dates of the population expansions were estimated from the data of both markers, and they suggest that these expansions happened between the Holocene and Middle Pleistocene epochs, which seem reasonable. For cytochrome *b*, the expansion of the left peak of the European lineage was not possible to estimate, because the result was zero. This was probably due to the inadequate resolution of cytochrome *b* to detect such recent events (Avise 2004). The intercalibration was done to obtain the rate of substitution for the HVII, and the time estimates were expected to result as nearly equal since the peaks were similar between the markers, but much older estimates were obtained for HVII. This suggests that the calibration was not completely successful, and it will be discussed thoroughly later. The estimates obtained for cytochrome *b* were then considered as more reliable and new estimates as close as possible to those of cytochrome *b* were calculated for HVII using a substitution rate of 12 %, and the results were 0.05 for the left peak, 0.20 for the right peak of the European lineage and 0.07 Mya for the peak of the Siberian lineage. These estimates are very inexact, thus no conclusions will be made about the climatic circumstances during the population expansion events.

At the final and smallest scale in Finland, the Northern Bat population was surprisingly different from that of the other countries from which HVII data was obtained: individuals with the Siberian lineage haplotypes were found only in Finland in contrast to the European lineage that inhabited

each of the sampled North European countries. The Finnish sample included only 17 Siberian lineage individuals as opposed to the 84 European lineage individuals (closely related individuals included in both data), thus, the Siberian lineage is rarer in Finland than the European lineage. Rarity of the Siberian lineage could result from a slow migration rate to Finland or from a change into European lineage due to mating of Siberian lineage males and European lineage females. This hypothesis is quite probable since Northern Bat males are the more dispersal sex of the species (Dietz *et al*. 2009).

The European lineage in Finland contains a lot of variation, similar in magnitude to that of the North European scale, but the variation in Finland was not geographically structured. The high diversity in the European lineage resulted from the clusters found at the North European scale. These clusters seem to have been formed before the recolonization of Fennoscandia, since individuals belonging to each cluster are quite evenly distributed in Finland. Furthermore, there are no obvious geographical boundaries in Finland that could have caused the formation of a subdivision to the present Finnish European lineage Northern Bat population. Thus, the HVII reflects the Middle and Late Pleistocene population structure already found from the European scale data.

The relatively low nucleotide and haplotype diversities of the Siberian lineage were already reported at the North European scale, and the sample size was too small and partly clustered so that testing for geographical structure in the genetic variation was not meaningful. However, by examining the sample location map (see Figure 3.6), the Siberian lineage can be seen to be clustered by most individuals to southeastern Finland. The concentration of the southeast resulted in larger amounts of samples collected in those areas. The composition of individuals in a village of Joroinen was interesting, since all the other colonies sampled in Finland consisted of individuals belonging only to the European lineage. In Joroinen, two closely located (~15 km apart) colonies were sampled in locations of Savuniemi and Kerisalonsaari. The colony of Kerisalonsaari comprised only individuals of the Siberian lineage but the colony of Savuniemi comprised individuals of both lineages, a most probable indication that the lineages intermix.

4.2 Refugia and post-glacial colonization routes

The northern Palearctic Northern Bats seem to have been occupying several different refugia during the glacials. The clusters I–V discovered from the North European Northern Bat HVII data

plausibly were formed as a result of population fragmentation into refugia during glacial advances. The clusters I–IV and V were in distinct geographical locations for so long that deeply diverged lineages, the European and Siberian respectively, were formed. The lineages likely retained separate from their divergence to the end of the last glaciation, since during glacials large areas would have been uninhabitable for bats, and the mountains and deserts of the Middle East reinforced the barrier (Hewitt 1996).

The Siberian lineage presumably inhabited eastern Russia, Kazakhstan, Mongolia and northernmost China as a single population during full-glacial conditions, since the lineage was found to form only one cluster. On the contrary, the European lineage consists of a few clusters that formed during glacials presumably in the three major European refugia, of which the Apennines Peninsula was likely occupied during glacials, since the contemporary range of the Northern Bat reaches to North Italy, south of the Alps. However, it does not reach the two other major refugia, the Balkans and Iberia (Wilson & Reeder 2005). Bennett *et al*. (1991) identified the western Balkan peninsula as the most important refugium for forests, so it is probable that Northern Bats occupied it during the Pleistocene and fossil findings could confirm their residence.

Serotine individuals of the West European lineage carrying the Northern Bat mitotypes existing in Iberia and Balkans (Figure 4.1) could indicate glacial residence of Northern Bats in those refugia, but this is not exclusionary evidence, since it is not certain how and where the introgression between the species took place. Juste *et al.* (2013) hypothesized that the West European Serotines captured Northern Bat mtDNA by migrating from Middle East to a new area in the West, which was already occupied by the resident Northern Bat. This Serotine migration resulted in a Northern Bat mitotype for all West European Serotines, but the scenario should be confirmed by further research. With additional HVII sequence data from the southern range it might be possible to deduce which Northern Bat cluster originated from the Apennines refugium, but the other two refugia not with total certainty. In *Myotis*, Bogdanowicz *et al*. (2009) found that Mouse-Eared Bats have apparently moved long distances (even from Iberia to the Balkans) so that within a refugium more than one haplogroups, which have originated in different refugia, are found. This might have happened in the Northern Bat also adding uncertainty to the evaluation of the source refugia of the clusters. Thus, as Taberlet *et al*. (1998) stated, "molecular genetics based on polymorphism present in extant organisms cannot by itself localize refugia with precision."

The estimation of post-glacial colonization routes from the mtDNA of extant individuals is also a challenge without comprehensive data from the southern areas of the Northern Bat's range, but possible to a certain degree by combining the available genetic information with geographical

locations of the individuals sampled. When an individual belonging to a certain cluster is found from a certain location, it is assumed that it is a "footprint" that the individuals of that cluster have left to the genomes of their progeny on the path the individuals of that cluster have passed. The path of recolonization of the Siberian lineage seems to have been quite simple, since this lineage only consisted of a single cluster, but the European lineage clusters presumably formed in more than one refugia also resulting to multiple colonization routes. A trend of reduction from south to north in genetic diversity is considered as a signature of migration from a source population (refugium) towards north (Hewitt 1996). In the Northern Bat European lineage the northwards reduction was not observed in the haplotype diversity values, but in the nucleotide diversity the trend was distinct. The correlation between genetic diversity and latitude has been detected also in other bat species, such as *Myotis myotis* (π, western colonies only; Ruedi & Castella 2003) and *Nyctalus noctula* (in haplotype diversity, but not π; Petit *et al*. 1999). In the Northern Bat the not very great, although significant, difference between the nucleotide diversity and non-existent difference in the haplotype diversity of latitudinally organized groups might have resulted from a relatively slow recolonization rate (Hewitt 1999). That multiple clusters from presumably different refugia are found in Fennoscandia also suggest a slow rate, thus, several clusters spread north simultaneously as opposed to one fast moving cluster (Hewitt 1996).

The results of this study suggest that the clusters of the European lineage arrived in Fennoscandia along at least three different pathways based on the pattern how individuals belonging to certain clusters are located in Figure 3.4: (i) The European lineage clusters II and IV both arrived via single routes from the south, (ii) clusters I and III by dichotomous routes also from the south, and (iii) the Siberian lineage cluster V from the east using a single path (Figure 4.1). The dichotomous routes of clusters I and III were: the Baltic countries, i.e. eastern route, and Denmark/Sweden, i.e. western route. It seems that these clusters originated from a single refugium, split prior to their arrival to Fennoscandia (Figure 4.1), based on the observation that in Latvia only these clusters were present, which suggests that they passed through that country, and individuals belonging to these clusters were also represented in all of the Fennoscandian countries (Finland, Sweden and Norway). Since clusters II and IV were not found in Latvia, they seem to have arrived to Fennoscandia only through the Denmark/Sweden pathway. The eastern and western routes have been recognized in other studies of different taxa, such as brown bear, oaks and newt *Triturus cristatus* (Hewitt 1999 and references therein). As already discussed, no conclusions can be drawn from which refugia each of the clusters originated from, since the Northern Bat's range does not reach the Balkans and Iberia and I have no data from the area of the Alps. Cluster V of the Siberian lineage was found only in

Finland (this study) and Russia (Artyushin *et al*. 2009; Kruskop *et al*. 2012) suggesting it is recolonizing Fennoscandia and perhaps central Europe from the east.

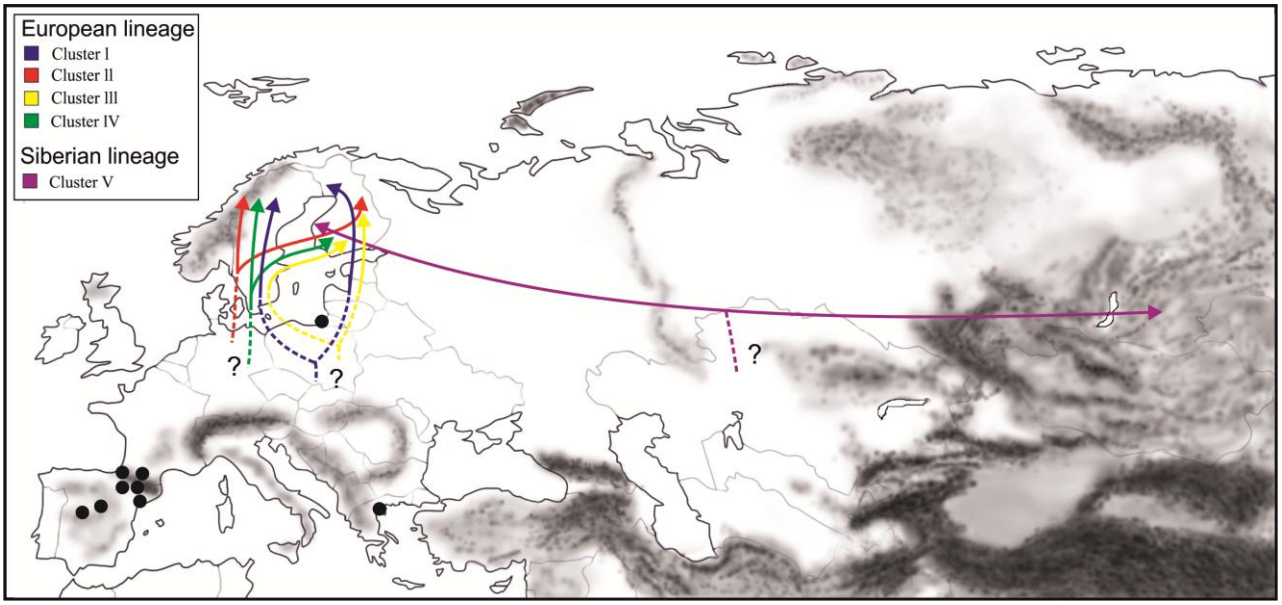


Figure 4.1 A scenario of the post-glacial colonization routes of the Northern Bat lineages. Color codes of the clusters refer to Figures 3.2 and 3.4. The arrow representing cluster V, the Siberian lineage, is based on data of Artyushin *et al*. (2009) and Kruskop *et al*. (2012) outside Finland. Dashed lines depict assumed routes; there are no data to confirm them. Locations of Serotine individuals of the West European lineage carrying the Northern Bat mitotype (after Artyushin *et al*. 2009) are marked as solid dots. Tinted areas illustrate mountains, and the elevation increases with darker color.

In further examination of the European lineage HVII data in Fennoscandia and Latvia, the values compared in pairwise Φ_{ST}'s differed most and were significant between Finland, Norway and Latvia, whereas the results from comparing these countries with Sweden were the lowest and non-significant, which could have resulted from two different scenarios: Either the European lineage clusters have been mixing in the area of Sweden after recolonization, or Sweden was colonized first and the other clusters were diverged from the Swedish population. The former scenario seems more probable, because in Sweden the within countries nucleotide and haplotype diversities were the highest of all sampled countries, and in many species previously allopatric populations arrived via both western and eastern recolonization routes and encountered in Sweden where the populations were characterized by high genetic diversity (e.g. Hewitt 1999). This encounter of different populations is called a contact zone (Provan & Bennett 2008). Contact zones in Sweden have been described, for instance, for Field Vole (*Microtus agrestis*; Jaarola & Tegelström 1995), Brown Bear (*Ursus arctos*; Taberlet & Bouvet 1994) and Moor Frog (*Rana arvalis*; Knopp & Merilä 2009). It must be brought to attention that when both the European and Siberian lineages are examined simultaneously, the highest genetic diversity is found in Finland where individuals representing each lineage are mixed indicating the secondary contact of the major maternal lineages.

The Northern Bat Siberian lineage seems to have colonized Fennoscandia, and particularly Finland, only from the east, Russia (Figure 4.1). It was not possible to examine if there is gradual decrease of genetic variation in the Siberian lineage, because the Russian Northern Bat data lacked HVII sequences and the sample size of additional cytochrome *b* sequences from Russia was very small. The Siberian lineage has not yet been observed in the west of Finland (this study) or far west from the western border of Russia in continental Europe, and the European lineage has not been observed in east further from the western border of Russia (Artyushin *et al*. 2009; Kruskop *et al*. 2012). The limited longitudinal spread of the lineages might result from elimination by competition when individuals of one lineage already occupied the space they reside and prevent the bats of the other lineage from spreading to new areas (Taberlet *et al*. 1998).

4.3 Evaluation of mitochondrial DNA markers used in this study

On a broad scale, with all or most of the data included, the two markers employed were appropriate tools for the study. The NJ tree topologies I obtained by both markers were consistent with those in previous studies, such as Ibáñez *et al*. (2006), Artyushin *et al*. (2009) and Kruskop *et al*. (2012), who, respectively, used the mitochondrial sequences of NADH dehydrogenase subunit I (ND1), cytochrome *b* and cytochrome oxidase subunit I (COI), and excluded the possibility of nuclear pseudogenes (Bensasson *et al*. 2001). However, some difficulties were encountered and will be discussed below.

From the mismatch distributions of the Northern Bat data, population expansion events were identified and the dates of these events were estimated for the two markers, HVII and cytochrome *b*. The different markers produced different estimates suggesting that the intercalibration was faulty. Cytochrome *b* and HVII evolve at different rates, and the assumed rate of cytochrome *b* (2 %/Myr) was used secondarily to assess the rate of HVII. The estimates for the Northern Bat population expansion times resulted with nearly twice as old estimates for HVII than for cytochrome *b*. Thus, an approximately 12 % HVII substitution rate would have been more suitable. In the noctule bat (*Nyctalus noctula*), Petit *et al*. (1999) estimated the rate of HVII as 6.3–25.2 %/Myr, which is even twice as much as the 12 % of the Northern Bat. The considerable difference between the cytochrome *b* and HVII time estimates in this study suggests that the regression did not measure the rate difference properly, since the line depicting dependence between the pairwise distance values of cytochrome *b* and HVII was linear, whereas it more likely might be curvilinear (see Figure 3.3).

In addition, the assumed rate of cytochrome *b* for the Northern Bat might also be inaccurate. Calibrations that were estimated using the fossil record for cytochrome *b* at least in two other vespertilionid bat groups resulted with a higher rate than the conventional 2 %/Myr. In long-eared bats (*Plecotus*) the rate was estimated as 3.5 % (Juste *et al*. 2004) and in the mouse-eared bats (*Myotis*) as 4.8 % (maximum likelihood corrected; Ruedi & Mayer 2001). Thus, the time estimate for the divergence of the Northern Bat lineages could be somewhat wrong. Although the genus *Eptesicus* has been examined quite extensively in the recent years (Artyushin *et al*. 2009; Kruskop *et al*. 2012; Juste *et al*. 2013), clock calibrations were not estimated for the mitochondrial markers used by these authors. Juste *et al*. (2013) studied the genus *Eptesicus* also by analyzing the nuclear DNA, and interestingly the introgression of the Northern Bat and Serotine did not show in the phylogenetic tree constructed of the data. A single marker (mtDNA in this case) by itself does not carry all the information of this species' history, and thus more evidence should be collected by using other independent molecular markers (Ballard & Whitlock 2004).

Choosing a correct marker is emphasized when considering the timescale under examination. A comprehensive sample of the Finnish Northern Bat population (mainly European lineage) was collected to examine possible population structure in that area, but the HVII marker failed to observe any structure that might have been formed after the colonization of Fennoscandia and reflected the ancient cluster structure that had formed during the Pleistocene epoch. This result confirmed that the variation in the Northern Bat European lineage was of ancient origin, but for studying the post-recolonization Finnish population a more highly variable marker, such as microsatellites, would have been more appropriate.

4.4 Prospects for future studies

The results of this study raised many further questions regarding the Northern Bat. It would be interesting to study the European and Russian populations further by mtDNA markers for obtaining a more comprehensive view of the European and Siberian lineage clusters in terms of refugia and the northward decrease of genetic diversity. Further sampling from central and north Sweden could also reveal if the Siberian lineage is spreading west. Comprehensive sampling from west Russia, close to the border of Finland and the Baltic countries, could confirm if the contact zone of the lineages is in East Europe close to the border of Russia (Artyushin *et al*. 2009; Kruskop *et al*. 2012) and reveal the width of the contact zone.

Additional studies of the Finnish Northern Bats could reveal if there are more colonies like Savuniemi that consist of both Northern Bat lineages, and if interbreeding is the reason why the Siberian lineage is so rare in Finland. Testing a marker mediated by Northern Bat males or both sexes would also be of great interest because these might give a different view of the species' history. This thesis studied the Northern Bats by a female-mediated marker, but the sexes of this species have different life styles: Female juveniles and adults return every summer to the colony they were born in, but, in autumn when nursery colonies disperse, juvenile males leave to live a solitary life and disperse more effectively (Dietz *et al*. 2009). Thus, from a male perspective the Northern Bat population could be structured differently or it could be a coherent single population.

5 CONCLUSIONS

The Northern Bat is a phylogeographically unique bat species. Currently, it has the northernmost as well as the most comprehensive distribution in Fennoscandia of all the North European bat species, and its population consists of two distinct mitochondrial lineages, which have formed between the Early and Middle Pleistocene epochs in separate geographical locations. Nowadays, the individuals of these lineages most likely interbreed at the locations of secondary contact and as characterizations of the bat's morphology seem not to be able to tell these lineages apart, no changes to the species nomenclature is needed. The genetic variation of the European lineage was found to be high, reflecting population division into separate glacial refugia. On the contrary, in the Siberian lineage genetic variation was relatively low and it formed a single cluster indicating that it remained as a coherent population during glaciations. Sweden seems to be a contact zone for the European lineage clusters, and Finland is the contact zone for the European and Siberian lineages. In general, the Pleistocene conditions had a substantial effect on the Northern Bats by splitting the population into smaller, temporarily isolated groups. Although the time estimates obtained here were not very accurate and are not completely reliable without further studies, the timescale, however, suggests that the events in the Northern Bat population history are ancient rather than recent. Furthermore, the history revealed here is from a genome, which has a female-mediated inheritance and thus male-restricted or biparental genes might show a different kind of history due to the different lifestyles of the two sexes. The two markers used in this study made it possible to examine of the Northern Bat population in slightly different resolutions, which enabled comparison between the markers and also preliminary characterization of the HVII for the future use on the Northern Bat.

6 ACKNOWLEDGEMENTS

I am grateful for my supervisors for making it possible for me to conduct this study on bat genetics. I will also express gratitude towards my better half, Janne, for being there for me, and towards my family and friends who have been involved with this work in anyway. Special thanks to Hanna Laakkonen, Katja Nylund, Laura Hekanaho and Miisa Virta respectively, for their help with the theory, in the lab, with the language and with everything else. I will also thank Lammi biological station for accommodation and amazing food during my two week stay, which was possible by the Gradut valmiiksi –scholarship. Suomen Biologian Seura Vanamo ry is thanked for the grant with which I participated in the XIIth European Bat Research Symposium held in Vilnius, Lithuania and Oskar Öflunds stiftelse for covering my laboratory expenses.

7 REFERENCES

Web resources:

Cohen, K. M., Finney, S. & Gibbard, P. L. (2013) Chronostratigraphic chart: http://www.stratigraphy.com

IUCN Red List of Threatened Species and bat range maps: http://www.iucnredlist.org; http://maps.iucnredlist.org

Maps: http://d-maps.com

Bibliography:

Aitken, M. J. & Stokes, S. (1997) Climatostratigraphy. *In*: Taylor, R. E. & Aitken, M. J. (Eds.) *Chronometric Dating in Archaeology*, New York, Plenum Press, pp. 1–30.

Aliabadian, M., Roselaar, C. S., Nijman, V., Sluys, R. & Vences, M. (2005) Identifying contact zone hotspots of passerine birds in the Palearctic region, *Biology Letters*, 1, 21–23.

Arbogast, B. S., Edwards, S. V., Wakeley, J., Beerli, P. & Slowinski, J. B. (2002) Estimating divergence times from molecular data on phylogenetic and population genetic timescales. *Annual Review of Ecology and Systematics*, 33, 707–740.

Artyushin, I. V., Bannikova, A. A., Lebedev, V. S. & Kruskop, S. V. (2009) Mitochondrial DNA relationships among North Palearctic *Eptesicus* (Vespertilionidae, Chiroptera) and past hybridization between Common Serotine and Northern Bat. *Zootaxa*, 2262, 40–52.

Avise, J. C. (1986) Mitochondrial DNA and the evolutionary genetics of higher animals. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences*, 312, 325–342.

Avise, J. C. (1989) Gene trees and organismal histories: A phylogenetic approach to population biology. *Evolution*, 43, 1192–1208.

Avise, J. C. (2000) *Phylogeography: The History and Formation of Species*. Harvard University Press, Cambridge, Massachusetts, 447 pp.

Avise, J. C. (2004) *Molecular Markers, Natural History and Evolution*. Sunderland, MA, Sinauer Associates, 684 pp.

Avise, J. C., Ball, R. M. & Arnold, J. (1988) Current versus historical population sizes in vertebrate species with high gene flow: A comparison based on mitochondrial lineages and inbreeding theory for neutral mutations. *Molecular Biology and Evolution*, 5, 331–344.

Avise, J. C., Walker, D. & Johns, G. C. (1998) Speciation durations and Pleistocene effects on vertebrate phylogeography. *Proceedings of the Royal Society London B*, 265, 1707–1712.

Avise, J. C., Bowen, B. W., Lamb, T., Meylan, A. B. & Bermingham, E. (1992) Mitochondrial DNA evolution at a turtle's pace: Evidence for low genetic variability and reduced microevolutionary rate in the *Testudines*. *Molecular Biology and Evolution*, 9, 457–473.

Avise, J. C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neiges, J. E., Reeb, C. A. & Saunders, N. C. (1987) Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. *Annual Reviews of Ecology, Evolution and Systematics*, 18, 489–522.

Ballard, J. W. O. & Whitlock, M. C. (2004) The incomplete natural history of mitochondria. *Molecular Ecology*, 13, 729–744.

Barrat, E. M., Deaville, R., Burland, T. M., Bruford, M. W., Jones, G., Racey, P. A. & Wayne, R. K. (1997) DNA answers to the call of *Pipistrelle* bat species. *Nature*, 387, 138–139.

Bennett, K. D., Tzedakis, P. C. & Willis, K. J. (1991) Quaternary refugia of north European trees. *Journal of Biogeography*, 18, 103–115.

Bensasson, D., Zhang, D.-X., Hartl, D. L. & Hewitt, G. M. (2001) Mitochondrial pseudogenes: evolution's misplaced witnesses. *Trends in Ecology and Evolution*, 16, 314–321.

Berthier, P., Excoffier, L. & Ruedi, M. (2006) Recurrent replacement of mtDNA and cryptic hybridization between two sibling bat species *Myotis myotis* and *Myotis blythii*. *Proceedings of the Royal Society London B, Biological Sciences*, 273, 3101–3109.

Bilton, D. T., Mirol, P. M., Mascheretti, S., Fredga, K., Zima, J. & Searle, J. B. (1998) Mediterranean Europe as an area of endemism for small mammals rather than source for northwards postglacial colonization. *Proceedings of the Royal Society London B*, 265, 1219–1229.

Birky, C. W. (2001) The inheritance of genes in mitochondria and chloroplasts: laws, mechanisms, and models. *Annual Review of Genetics*, 35, 125–148.

Birky, C. W. Jr. (2005) Uniparental inheritance of mitochondrial and chloroplast genes: Mechanisms and evolution. *Proceedings of the National Academy of Sciences of the United States of America*, 92, 11331–11338.

Bogdanowicz, W., Van Den Bussche, R. A., Gajewska, M., Postawa, T. & Harutyunyan, M. (2009) Ancient and contemporary DNA sheds light on the history of Mouse-eared bats in Europe and the Caucasus. *Acta Chiropterologica*, 11, 289–305.

Bradley, R. D. & Baker, R. J. (2001) A test of the genetic species concept: Cytochrome-b sequences and mammals. *Journal of Mammalogy*, 82, 960–973.

Bromham, L. & Penny, D. (2003) The modern molecular clock. *Nature Reviews Genetics*, 4, 216–224.

Castella, V., Ruedi, M. & Excoffier, L. (2001) Contrasted patterns of mitochondrial and nuclear structure among nursery colonies of the bat Myotis myotis. *Journal of Evolutionary Biology*, 14, 708–720.

Clement, M., Posada, D. & Crandall, K. A. (2000) TCS: A computer program to estimate gene genealogies. *Molecular Ecology*, 9, 1657–1659.

Constantine, D. G. (1958) An automatic bat-collecting device. *Journal of Wildlife Management*, 22, 17–22.

Crochet, P. A. & Desmarais, E. (2000) Slow rate of evolution in the mitochondrial control region of gulls (Aves: Laridae). *Molecular Biology and Evolution*, 17, 1797–1806.

Dietz, C., von Helversen, O. & Nill, D. (2009) *Bats of Britain, Europe & Northwest Africa*. London, A & C Black Publishers Ltd, 400 pp.

Drummond, A. J., Nichills, G. K., Rodrigo, A. G. & Solomon, W. (2002) Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. *Genetics*, 161, 1307–1320.

Excoffier, L., Laval, G. & Schneider, S. (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, 1, 47–50.

Excoffier, L., Smouse, P. E. & Quattro, J. M. (1992) Analysis of Molecular Variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics*, 131, 479–491.

Esposti, M. D., De Vries, S., Crimi, M., Ghelli, A., Patarnello, T. & Meyer, A. (1993) Mitochondrial cytochrome b: evolution and structure of the protein. *Biocimica et Biophysica Acta*, 1143, 243–271.

Fleischer, R. C., McIntosh, C. E. & Tarr, C. L. (1998) Evolution on a volcanic conveyor belt: using phylogeographic reconstructions and K-Ar based ages of the Hawaiian Islands to estimate molecular evolutionary rates. *Molecular Ecology*, 7, 533–545.

Goodman, S. M., Maminirina, C. P., Bradman, H. M., Christidis, L. & Appleton, B. (2009) The use of molecular phylogenetic and morphological tools to identify cryptic and paraphyletic species: Examples from the diminutive long-fingered bats (Chiroptera: Miniopteridae: *Miniopterus*) on Madagascar. *American Museum Novitates*, 3669, 1–34.

Hall, T. A. (1999) BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, 41, 95–98.

Hamilton, M. B. (2009) *Population Genetics*. Wiley-Blackwell, Chichester, West Sussex, 424 pp.

Hewitt, G. M. (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, 58, 247–276.

Hewitt, G. M. (1999) Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, 68, 87–112.

Hill, D. A. & Greenaway, F. (2005) Effectiveness of an acoustic lure for surveying bats in British woodlands. *Mammal Review*, 35, 116–122.

Hoffman, F. G., Owen, J. G. & Baker, R. J. (2003) mtDNA perspective of chromosomal diversification and hybridization in Peter's tent-making bat (*Uroderma bilobatum*: Phyllostomidae). *Molecular Ecology*, 12, 2981–2993.

Howell, N. (1989) Evolutionary conservation of protein regions in the proton-motive cytochrome *b* and their possible roles in redox catalysis. *Journal of Molecular Evolution*, 29, 157–169.

Hulva, P., Horáček, I., Strelkov, P. & Benda, P. (2004) Molecular architecture of *Pipistrellus pygmaeus* complex (Chiroptera: Vespertilionidae): further cryptic species and Mediterranean origin of the divergence. *Molecular Phylogenetics and Evolution*, 32, 1023–1035.

Huntley, B. & Webb, T., III. (1989) Migration: species' response to climatic variations caused by changes in the earth's orbit. *Journal of Biogeography*, 16, 5–19.

Hutcheon, J. M. & Kirsch, J. A. W. (2006) A moveable face: deconstructing the Microchiroptera and a new classification of extant bats. *Acta Chiropterologica*, 8, 1–10.

Hutchison, C. A. III, Newbold, J. E., Potter, S. S. & Edgell, M. H. (1974) Maternal inheritance of mammalian mitochondrial DNA. *Nature*, 251, 536–538.

Ibáñez, C., Garcia-Mudarra, J. L., Ruedi, M., Stadelmann, B. & Juste, J. (2006) The Iberian contribution to cryptic diversity in European bats. *Acta Chiropterologica*, 8, 277–297.

Irwin, D. M., Kocher, T. D. & Wilson, A. C. (1991) Evolution of the cytochrome b gene of mammals. *Journal of Molecular Evolution*, 32, 128–144.

Jaarola, M. & Tegelström, H. (1995) Colonization history of north European field voles (*Microtus agrestis*) revealed by mitochondrial DNA. *Molecular Ecology*, 4, 299–310.

Jukes, T. H. & Cantor, C. R. (1969) Evolution of protein molecules. *In*: Munro, H. (Ed.) *Mammalian protein metabolism*, Academic Press, New York, pp. 21–132.

Juste, J., Benda, P., Garcia-Mudarra, J. L. & Ibáñez, C. (2013) Phylogeny and systematics of Old World serotine bats (genus *Eptesicus*, Verspertilionidae, Chiroptera): an integrative approach. *Zoologica Scripta*, doi: 10.1111/zsc.12020.

Juste, J., Ibáñez, C., Muñoz, J., Trujillo, D., Benda, P., Karataş, A. & Ruedi, M. (2004) Mitochondrial phylogeography of the long-eared bats (*Plecotus*) in the Mediterranean Palearctic and Atlantic Islands. *Molecular Phylogenetics and Evolution*, 31, 1114 –1126.

Kimura, M. (1968) Evolutionary rate at the molecular level. *Nature*, 217, 624–626.

Knopp, T. & Merilä, J. (2009) The postglacial recolonization of Northern Europe by *Rana arvalis* as revealed by microsatellite and mitochondrial DNA analysis. *Heredity*, 102, 174–181.

Kruskop, S. V., Borisenko, A. V., Ivanova, N. V., Lim, B. K. & Eger, J. L. (2012) Genetic diversity of northeastern Palaearctic bats as revealed by DNA barcodes. *Acta Chiropterologica*, 14, 1–14.

Kvist, L., Ruokonen, M., Lumme, J. & Orell, M. (1999) The colonization history and present-day population structure of the European great tit (*Parus major major*). *Heredity*, 82, 495–502.

Librado, P. & Rozas, J. (2009) DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, 25, 1451–1452.

Liukko, U.-M., Henttonen, H., Hanski, I. K., Kauhala, K., Kojola, I. & Kyheröinen E.-M. (2010): Nisäkkäät. *In*: Rassi, P., Hyvärinen, E., Juslén, A. & Mannerkoski, I. (Eds.) (2010): Suomen lajien uhanalaisuus. Punainen kirja 2010./The 2010 Red List of Finnish Species. Ympäristöministeriö & Suomen ympäristökeskus, Helsinki, pp. 311–319.

Lutz, S., Wittig, H., Weisser, H.-J., Heizmann, J., Junge, A., Dimo-Simonin, N., Parson, W., Edelmann, J., Anslinger, K., Jung, S. & Augustin, C. (2000) Is it possible to differentiate mtDNA by means of HVIII in samples that cannot be distinguished by sequencing the HVI and HVII regions? *Forensic Science International*, 113, 97–101.

Machado, C. A. & Hey, J. (2003) The causes of phylogenetic conflict in a classic Drosophila species group. *Proceedings of the Royal Society B*, 270, 1193–1202.

Mao, X. G., Zhu, G. J., Zhang, S. & Rossiter, S. J. (2010) Pleistocene climatic cycling drives intra-specific diversification in the intermediate horseshoe bat (*Rhinolophus affinis*) in Southern China. *Molecular Ecology*, 19, 2754–2769.

Mao, X., Thong, V. D., Bates, P. J. J., Jones, G., Zhang, S. & Rossiter, S. J. (2013) Multiple cases of asymmetric introgression among horseshoe bats detected by phylogenetic conflicts across loci. *Biological Journal of the Linnean Society*, 110, 346–361.

Mayer, F. & von Helversen, O. (2001) Cryptic diversity in European bats. *Proceedings of the Royal Society B*, 268, 1825–1832.

Mayer, F., Dietz, C. & Kiefer, A. (2007) Molecular species identification boosts bat diversity. *Frontiers in Zoology* 4, doi:10.1186/1742-9994-4-4.

Meganathan, P., Pagan, H., McCulloch, E., Stevens, R. & Ray, D. (2012) Complete mitochondrial genome of three bat species and whole genome mitochondrial analyses reveal patterns of codon bias and lend support to a basal split in Chiroptera. *Gene*, 492, 121–129.

Meyer, A. (1994) Shortcomings of the cytochrome b gene as a molecular marker. *Trends in Ecology and Evolution*, 9, 209–214.

Nei, M. (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York, 512 pp.

Nei, M. & Kumar, S. (2000) *Molecular Evolution and Phylogenetics*. Oxford University Press, New York, 352 pp.

Pesole, G., Gissi, C., De Chirico, A. & Saccone, C. (1999) Nucleotide substitution rate of mammalian mitochondrial genomes. *Journal of Molecular Evolution*, 48, 42–434.

Petit, E., Excoffier, L. & Mayer, F. (1999) No evidence of bottleneck in the postglacial recolonization of Europe by the noctule bat (*Nyctalus noctula*). *Evolution*, 53, 1247–1258.

Petri, B., Von Haeseler, A. & Pääbo, S. (1996) Extreme sequence heteroplasmy in bat mitochondrial DNA. *Biological Chemistry*, 377, 661–667.

Provan, J. & Bennett, K. D. (2008) Phylogeographic insights into cryptic glacial refugia. *Trends in Ecology and Evolution*, 23, 564–571.

Pääbo, S., Poinar, H., Serre, D., Jaenicke-Després, V., Hebler, J., Rohland, N., Kuch, M., Krause, J., Vigilant, L. & Hofreiter, M. (2004) Genetic analyses from ancient DNA. *Annual Review of Genetics*, 38, 645–679.

Rabinowitz, M. & Swift, H. (1970) Mitochondrial nucleic acids and their relation to the biogenesis of mitochondria. *Physiological Reviews*, 50, 376–427.

Randi, E. (2000) Mitochondrial DNA. *In*: Baker, A. J. (Ed.) *Molecular Methods in Ecology*, Blackwell, London, pp. 136–167.

Rebelo, H., Froufe, E., Brito, J. C., Russo, D., Cistrone, L., Ferrand, N. & Jones, G. (2012) Postglacial colonization of Europe by the barbastelle bat: agreement between molecular data and past predictive modeling. *Molecular Ecology*, 21, 2761–2774.

Rodríguez, V., Brown, R. P., Terrasa, B., Pérez-Mellado, V., Castro. J. A., Picornell, A. & Ramon, M. M. (2013) Multilocus genetic diversity and historical biogeography of the endemic wall lizard from Ibiza and Formentera, *Podarcis pityusensis* (Squamata: Lacertidae). *Molecular Ecology*, 22, 4829–4841.

Rogers, A. R. & Harpending, H. (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution*, 9, 552–569.

Ruedi, M. & Mayer, F. (2001) Molecular systematic of bats of the genus *Myotis* (Vespertilionidae) suggests deterministic ecomorphological convergences. *Molecular Phylogenetics and Evolution*, 21, 436–448.

Ruedi, M. & Castella, V. (2003) Genetic consequences of the ice ages on nurseries of the bat *Myotis myotis*: a mitochondrial and nuclear survey. *Molecular Ecology*, 12, 1527–1540.

Rydell, J. (1989) Occurence of bats in northernmost Sweden (65˚N) and their feeding ecology in summer. *Journal of Zoology*, 227, 517–529.

Saitou, N. & Nei, M. (1987) The Neighbor-Joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, 4, 406–425.

Scheffler, I. E. (1999) *Mitochondria*. New York, Chichester, Wiley-Liss, 351 pp.

Siivonen, Y. & Wermundsen, T. (2008) Characteristics of winter roosts of bat species in southern Finland. *Mammalia*, 72, 50–56.

Stubbe, M., Ariunbold, J., Buuveibaatar, V., Dorjderem, S., Monkhzul, Ts., Otgonbaatar, M., Tsogbadrakh, M., Hutson, A.M., Spitzenberger, F., Aulagnier, S. & Coroiu, I. (2008) *Eptesicus nilssonii*. *In*: IUCN 2011. IUCN Red List of Threatened Species. Version 2011.2. URL: [www.iucnredlist.org]. Downloaded on **11 April 2012**.

Sun, K., Feng, J., Jin, L., Liu, Y., Shi, L. & Jiang, T. (2009) Structure, DNA sequence variation and phylogenetic implications of the mitochondrial control region in horseshoe bats. *Mammalian Biology*, 74, 130–144.

Taberlet, P. & Bouvet, J. (1994) Mitochondrial DNA polymorphism, phylogeography, and conservation genetics of the brown bear (*Ursus arctos*) in Europe. *Proceedings of the Royal Society of London B*, 255, 195–200.

Taberlet, P., Swenson, J. E., Sandengren, F. & Bjärvall, A. (1995) Localization of a contact zone between two highly divergent mitochondrial DNA lineages of the brown bear (*Ursus arctos*) in Scandinavia. *Conservation Biology*, 9, 1255–1261.

Taberlet, P., Fumagalli, L., Wust-Saucy, A.-G. & Cosson, J.-F. (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, 7, 453–464.

Tamura, K., Peterson, D., Stecher, G., Nei, M. & Kumar, S. (2011) MEGA 5: Molecular evolutionary genetics analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony methods. *Molecular Biology and Evolution*, 28, 2731–2739.

Teeling, E. C., Scally, M., Kao, D. J., Romagnoli, M. L., Springer, M. S. & Stanhope, M. J. (2000) Molecular evidence regarding the origin of echolocation and flight in bats. *Nature*, 403, 188–192.

Tegelström, H. (1987) Transfer of mitochondrial DNA from the northern red-backed vole (*Clethrionomys rutilus*) to the bank vole (*C. glareolus*). *Journal of Molecular Evolution*, 24, 218–227.

Tuttle, M. D. (1974) An improved trap for bats. *Journal of Mammalogy*, 55, 475–477.

Vallo, P., Benda, P., Cerveny, J. & Koubek, P. (2012) Conflicting mitochondrial and nuclear paraphyly in small-sized West African house bats (Vespertilionidae). *Zoologica Scripta*, 42, 1–12.

Webb, T. III & Bartlein, P. J. (1992) Global changes during the last 3 million years: climatic controls and biotic response. *Annual Review of Ecology and Systematics*, 23, 141–173.

Webb, N. J. & Tidemann, C. R. (1995) Hybridization between black (*Pteropus alecto*) and grey-headed (*P. poliocephalus*) flying foxes (Megachiroptera: Pteropodidae). *Australian Mammalogy*, 18, 63–72.

Wilkinson, G. S., Mayer, F., Kerth, G. & Petri, B. (1997) Evolution of repeated sequence arrays in the D-loop region of bat mitochondrial DNA. *Genetics*, 146, 1035–1048.

Wilmer, W. J., Moritz, C., Hall, L. & Toop, J. (1994) Extreme population structuring in the threatened ghost bat, Macroderma gigas: evidence from mitochondrial DNA. *Proceedings of the Royal Society London B*, 257, 193–198.

Wilson, C. E. & Reeder, D. M. (2005) *Mammal Species of the World: A Taxonomic and Geographic Reference*. Baltimore, MD, John Hopkins University Press, 2142 pp.

Worthington Wilmer, J. & Barratt, E. (1996) A non-lethal method of tissue sampling for genetic studies of chiropterans. *Bat Research News*, 37, 1–3.

Xu, L., He, C., Shen, C., Jiang, T., Shi, L., Sun, K., Berquist, S. W. & Feng, J. (2010) Phylogeography and population genetic structure of the great leaf-nosed bat (*Hipposideros armiger*) in China. *Journal of Heredity*, 101, 562–572.

Zuckerkandl, E. & Pauling. L. (1965) Evolutionary divergence and convergence in proteins. *In*: Bryson, V. & Vogel, H. J. (Eds.) *Evolving genes and proteins*, Academic Press, New York, pp. 97–166.

Photographs:

All photographs were taken by me.

8 APPENDICES

APPENDIX A

List of Northern Bat (*Eptesicus nilssonii*) samples (sorted by country: Fin, Finland; Lat, Latvia; Nor, Norway; Swe, Sweden), collection locations, haplotypes, museums (MZH, Finnish Museum of Natural History; ZMUO, Zoological Museum of the University of Oulu; LUA, Latvia University of Agriculture; RMZ, Helgeland Museum; NRM, Swedish Museum of Natural History), markers (cyt *b*, cytochrome *b*; HVII, hypervariable segment II), mtDNA lineages (EUR, European lineage; SIB, Siberian lineage) and latitudes of the sample locations for the samples used in this study.

Sample code	Location	Haplotype	Museum	Marker	Lineage	Latitude
Fin001	Åland	HP2/HP56	MZH	Cyt *b*/HVII	EUR	60.254986
Fin002	Åland	HP70	MZH	HVII	EUR	60.139545
Fin003	Åland	HP64	MZH	HVII	EUR	60.17508
Fin004	Åland	HP28	MZH	HVII	EUR	60.229926
Fin005	Åland	HP47	MZH	HVII	EUR	60.17508
Fin006	Åland	HP45	MZH	HVII	EUR	60.17508
Fin007	Vantaa	HP1/HP45	MZH	Cyt *b*/HVII	EUR	60.342883
Fin008	Porvoo	HP13/HP72	MZH	Cyt *b*/HVII	SIB	
Fin009	Ruokolahti	HP3/HP47	MZH	Cyt *b*/HVII	EUR	61.2863
Fin010	Dragsfjärd	HP2/HP39	MZH	Cyt *b*/HVII	EUR	60.021297
Fin011	Hanko	HP8/HP63	MZH	Cyt *b*/HVII	EUR	59.828346
Fin012	Urjala	HP8/HP62	MZH	Cyt *b*/HVII	EUR	61.080306
Fin013	Miehikkälä	HP6/HP36	MZH	Cyt *b*/HVII	EUR	60.652024
Fin014	Asikkala	HP8/HP33	MZH	Cyt *b*/HVII	EUR	61.29915
Fin015	Savitaipale	HP8/HP66	MZH	Cyt *b*/HVII	EUR	61.197533
Fin016	Eura	HP8/HP64	MZH	Cyt *b*/HVII	EUR	61.12958
Fin017	Myrskylä	HP2/HP47	MZH	Cyt *b*/HVII	EUR	60.669722
Fin018	Jämsä	HP5/HP51	MZH	Cyt *b*/HVII	EUR	61.86242
Fin019	Västanfjärd	HP2/HP46	MZH	Cyt *b*/HVII	EUR	60.049948
Fin020	Tampere	HP8/HP66	MZH	Cyt *b*/HVII	EUR	61.470222
Fin021	Anjalankoski	HP3	MZH	Cyt *b*	EUR	
Fin022	Ruovesi	HP8/HP16	MZH	Cyt *b*/HVII	EUR	61.951997
Fin023	Hankasalmi	HP7/HP24	MZH	Cyt *b*/HVII	EUR	62.389441
Fin024	Renko	HP8/HP16	MZH	Cyt *b*/HVII	EUR	60.893689
Fin025	Ilomantsi	HP64	MZH	HVII	EUR	62.667885
Fin026	Lappajärvi	HP8/HP64	MZH	Cyt *b*/HVII	EUR	63.252577
Fin027	Rantasalmi	HP3/HP48	MZH	Cyt *b*/HVII	EUR	62.038334
Fin028	Joensuu	HP3/HP47	MZH	Cyt *b*/HVII	EUR	62.60109
Fin029	Nilsiä	HP4/HP44	MZH	Cyt *b*/HVII	EUR	63.050566
Fin030	Pietarsaari	HP49	MZH	HVII	EUR	63.674244
Fin031	Vyborg	HP74	MZH	HVII	SIB	
Fin032	Kristinestad	HP8/HP29	MZH	Cyt *b*/HVII	EUR	62.273935
Fin033	Hogland	HP42	MZH	HVII	EUR	60.056884

Appendix A continues.

Sample code	Location	Haplotype	Museum	Marker	Lineage	Latitude
Fin026	Lappajärvi	HP8/HP64	MZH	Cyt *b*/HVII	EUR	63.252577
Fin027	Rantasalmi	HP3/HP48	MZH	Cyt *b*/HVII	EUR	62.038334
Fin028	Joensuu	HP3/HP47	MZH	Cyt *b*/HVII	EUR	62.60109
Fin029	Nilsiä	HP4/HP44	MZH	Cyt *b*/HVII	EUR	63.050566
Fin030	Pietarsaari	HP49	MZH	HVII	EUR	63.674244
Fin031	Vyborg	HP74	MZH	HVII	SIB	
Fin032	Kristinestad	HP8/HP29	MZH	Cyt *b*/HVII	EUR	62.273935
Fin033	Hogland	HP42	MZH	HVII	EUR	60.056884
Fin034	Hirvensalmi	HP3/ HP47	MZH	Cyt *b*/HVII	EUR	61.604012
Fin035	Hanko	HP8/HP16	MZH	Cyt *b*/HVII	EUR	59.828346
Fin036	Varkaus	HP59	MZH	HVII	EUR	62.31423
Fin037	Salo	HP53	MZH	HVII	EUR	60.344516
Fin038	Nauvo	HP20	MZH	HVII	EUR	60.148389
Fin039	Nauvo	HP2/HP46	MZH	Cyt *b*/HVII	EUR	60.192716
Fin040	Rauma	HP14/HP78	MZH	Cyt *b*/HVII	SIB	
Fin041[a]	Savuniemi	HP13/HP71	MZH	Cyt *b*/HVII	SIB	
Fin042[b]	Kerisalonsaari	HP12 (cyt *b*)	MZH	Cyt *b*/HVII	SIB	
Fin043[a]	Savuniemi	HP2/HP47		Cyt *b*/HVII	EUR	62.16314
Fin044[b]	Kerisalonsaari			Cyt *b*/HVII	SIB	
Fin045[a]	Savuniemi	HP13/HP79		Cyt *b*/HVII	SIB	
Fin046[b]	Kerisalonsaari			Cyt *b*/HVII	SIB	
Fin047[b]	Kerisalonsaari			Cyt *b*/HVII	SIB	
Fin048[b]	Kerisalonsaari			Cyt *b*/HVII	SIB	
Fin049[a]	Savuniemi	HP13/HP74		Cyt *b*/HVII	SIB	
Fin050[a]	Savuniemi			Cyt *b*/HVII	SIB	
Fin051[a]	Savuniemi			Cyt *b*/HVII	SIB	
Fin052[b]	Kerisalonsaari	HP73 (HVII)		Cyt *b*/HVII	SIB	
Fin053	Espoo	HP6/HP36		Cyt *b*/HVII	EUR	60.13594
Fin054	Kirkkonummi	HP8/HP23		Cyt *b*/HVII	EUR	59.987545
Fin055	Haukipudas	HP16	ZMUO	HVII	EUR	65.176281
Fin056	Ylikiiminki	HP2/HP50	ZMUO	Cyt *b*/HVII	EUR	65.028614
Fin057	Haukipudas	HP25	ZMUO	HVII	EUR	65.176281
Fin058	Haukipudas	HP16	ZMUO	HVII	EUR	65.176281
Fin059	Muhos	HP40	ZMUO	HVII	EUR	64.807538
Fin060	Tervo	HP14/HP76	ZMUO	Cyt *b*/HVII	SIB	
Fin061	Simo	HP25	ZMUO	HVII	EUR	65.658365
Fin062	Kiuruvesi	HP47	ZMUO	HVII	EUR	63.653802
Fin063	Muhos	HP49	ZMUO	HVII	EUR	64.790544
Fin064	Oulu	HP64	ZMUO	HVII	EUR	65.027246
Fin065	Pudasjärvi	HP47	ZMUO	HVII	EUR	65.361853
Fin066	Kajaani	HP77	ZMUO	HVII	SIB	
Fin067	Oulu	HP41	ZMUO	HVII	EUR	64.984005

Appendix A continues.

Sample code	Location	Haplotype	Museum	Marker	Lineage	Latitude
Fin068	Kuusamo	HP3/HP47	ZMUO	Cyt *b*/HVII	EUR	65.964567
Fin069	Kiiminki	HP33	ZMUO	HVII	EUR	65.148446
Fin070	Kiiminki	HP47	ZMUO	HVII	EUR	65.148446
Fin071	Kiiminki		ZMUO	HVII	EUR	
Fin072	Rauma	HP26	ZMUO	HVII	EUR	61.202091
Fin073	Pudasjärvi	HP21	ZMUO	HVII	EUR	65.361853
Fin074	Oulu	HP8/HP62	ZMUO	Cyt *b*/HVII	EUR	65.012615
Fin075[d]	Siuntio	HP8/HP27		Cyt *b*/HVII	EUR	60.097271
Fin076[c]	Hamina	HP3/HP43		Cyt *b*/HVII	EUR	60.686912
Fin077[c]	Hamina			Cyt *b*/HVII	EUR	
Fin078[d]	Siuntio			Cyt *b*/HVII	EUR	
Fin079[c]	Hamina			Cyt *b*/HVII	EUR	
Fin080[c]	Hamina	HP47		Cyt *b*/HVII	EUR	60.686912
Fin081[c]	Hamina			HVII	EUR	
Fin082[c]	Hamina			Cyt *b*/HVII	EUR	
Fin083[c]	Hamina			HVII	EUR	
Fin084[c]	Hamina			HVII	EUR	
Fin085[c]	Hamina			Cyt *b*/HVII	EUR	
Fin086[c]	Hamina			Cyt *b*/HVII	EUR	
Fin087[e]	Loviisa	HP8/HP31		Cyt *b*/HVII	EUR	60.392988
Fin088[e]	Loviisa			Cyt *b*/HVII	EUR	
Fin089	Pyhtää	HP34 (HVII)		Cyt *b*/HVII	EUR	60.500248
Fin090	Inkoo	HP37		Cyt *b*/HVII	EUR	60.01731
Fin091[d]	Siuntio	HP21 (HVII)		Cyt *b*/HVII	EUR	60.097271
Fin092[d]	Siuntio			Cyt *b*/HVII	EUR	
Fin093[d]	Siuntio	HP21 (HVII)		Cyt *b*/HVII	EUR	60.097271
Fin094[d]	Siuntio	HP16 (HVII)		Cyt *b*/HVII	EUR	60.097271
Fin095[a]	Savuniemi			Cyt *b*/HVII	EUR	
Fin096[a]	Savuniemi	HP75 (HVII)		Cyt *b*/HVII	SIB	
Fin097[c]	Hamina			Cyt *b*/HVII	EUR	
Fin098[a]	Savuniemi			Cyt *b*/HVII	EUR	
Fin099[e]	Loviisa			Cyt *b*/HVII	EUR	
Fin100[e]	Loviisa			HVII	EUR	
Fin101[e]	Loviisa			Cyt *b*/HVII	EUR	
Fin102[e]	Loviisa			Cyt *b*/HVII	EUR	
Lat01	Madonas	HP18	LUA	HVII	EUR	56.854407
Lat02	Madonas	HP8/HP66	LUA	Cyt *b*/HVII	EUR	56.854407
Lat03	Jelgava	HP8/HP18	LUA	Cyt *b*/HVII	EUR	56.652608
Lat04	Riga	HP17	LUA	HVII	EUR	56.962629
Lat05	Tervete	HP8/HP65	LUA	Cyt *b*/HVII	EUR	56.480065
Lat06	Amatas	HP8/HP66	LUA	Cyt *b*/HVII	EUR	57.318993
Lat07	Jelgava	HP34	LUA	HVII	EUR	56.652608

Appendix A continues.

Sample code	Location	Haplotype	Museum	Marker	Lineage	Latitude
Lat08	Ligatne	HP20	LUA	HVII	EUR	57.2358
Lat09	Ligatne	HP30	LUA	HVII	EUR	57.2358
Lat10	Incukalna	HP34	LUA	HVII	EUR	56.994708
Lat11	Ligatne	HP16	LUA	HVII	EUR	57.2358
Lat12	Jelgava	HP18	LUA	HVII	EUR	56.652608
Lat13	Riga	HP34	LUA	HVII	EUR	56.946203
Nor01	Rana	HP2/HP58	RMZ	Cyt *b*/HVII	EUR	66.320535
Nor02	Hemnes	HP2/HP52	RMZ	Cyt *b*/HVII	EUR	66.00627
Nor03	Hemnes	HP11/HP24	RMZ	Cyt *b*/HVII	EUR	66.070135
Nor04	Rana	HP52	RMZ	HVII	EUR	66.315786
Nor05	Rana	HP22	RMZ	HVII	EUR	66.375613
Nor06	Hemnes	HP47	RMZ	HVII	EUR	66.193717
Nor07	Rana	HP16	RMZ	HVII	EUR	66.312485
Swe01	Skåne	HP8/HP32	NRM	Cyt *b*/HVII	EUR	55.70466
Swe02	Skåne	HP8/HP38	NRM	Cyt *b*/HVII	EUR	55.990257
Swe03	Uppland	HP3/HP60	NRM	Cyt *b*/HVII	EUR	59.314449
Swe04	Uppland	HP8/HP16	NRM	Cyt *b*/HVII	EUR	59.527238
Swe05	Bohuslän	HP8/HP61	NRM	Cyt *b*/HVII	EUR	58.349415
Swe06	Uppland	HP8/HP19	NRM	Cyt *b*/HVII	EUR	59.480277
Swe07	Uppland	HP16	NRM	HVII	EUR	60.345896
Swe08	Halland	HP8/HP23	NRM	Cyt *b*/HVII	EUR	56.674375
Swe09	Gotland	HP9/HP68	NRM	Cyt *b*/HVII	EUR	57.833777
Swe10	Medelpad	HP8/HP32	NRM	Cyt *b*/HVII	EUR	62.390811
Swe11	Småland	HP8/HP67	NRM	Cyt *b*/HVII	EUR	57.174712
Swe12	Uppland	HP2/HP55	NRM	Cyt *b*/HVII	EUR	60.194976
Swe13	Södermanland	HP34	NRM	HVII	EUR	59.242595
Swe14	Gotland	HP9/HP69	NRM	Cyt *b*/HVII	EUR	57.833777
Swe15	Gästrikland	HP8/HP35	NRM	Cyt *b*/HVII	EUR	60.560666
Swe16	Uppland	HP16	NRM	HVII	EUR	59.434941
Swe17	Bohuslän	HP47	NRM	HVII	EUR	57.893791
Swe18	Västergötland	HP54	NRM	HVII	EUR	58.500415
Swe19	Östergötland	HP2/HP55	NRM	Cyt *b*/HVII	EUR	58.346037
Swe20	Södermanland	HP8/HP16	NRM	Cyt *b*/HVII	EUR	58.968196
Swe21	Gotland	HP15	NRM	HVII	EUR	57.15376
Swe22	Hälsingland	HP16	NRM	HVII	EUR	61.642823
Swe23	Skåne	HP57	NRM	HVII	EUR	55.48045
Swe24	Småland	HP58	NRM	HVII	EUR	56.568675

[a–e] Colony locations: a – Savuniemi; Joroinen; b – Kerisalonsaari, Joroinen; c – Hamina; d – Siuntio; e – Loviisa.

APPENDIX B

List of species (*E., Eptesicus*), sample locations, haplotypes (and marker; cyt *b*, cytochrome *b*), mtDNA lineages (EUR, European lineage; SIB, Siberian lineage; W EUR, West European lineage; ORI, Original lineage) and GenBank accession numbers for additional cytochrome *b* sequences retrieved from GenBank.

Species	Location	Haplotype	Lineage	GenBank Number	Source
E. nilssonii	Germany	HP10 (Cyt *b*)	EUR	AF37683	Ruedi & Mayer 2001
E. nilssonii	Bryansk	HP8 (Cyt *b*)	EUR	GQ272565	Artyushin *et al*. 2009
E. nilssonii	Chita	HP14 (Cyt *b*)	SIB	GQ272569	Artyushin *et al*. 2009
E. nilssonii	Chita	HP14 (Cyt *b*)	SIB	GQ272568	Artyushin *et al*. 2009
E. nilssonii	Yenisey	HP13 (Cyt *b*)	SIB	GQ272582	Artyushin *et al*. 2009
E. nilssonii	Yenisey	HP14 (Cyt *b*)	SIB	GQ272567	Artyushin *et al*. 2009
E. serotinus	Bryansk		ORI	GQ272566	Artyushin *et al*. 2009
E. serotinus	Greece		W EUR	AF376837	Ruedi & Mayer 2001
E. serotinus	Kaliningrad		W EUR	GQ272587	Artyushin *et al*. 2009
E. serotinus	Kabardino-Balkaria		ORI	GQ272578	Artyushin *et al*. 2009
E. serotinus	Krasnodar		ORI	GQ272575	Artyushin *et al*. 2009
E. serotinus	Krasnodar		ORI	GQ272571	Artyushin *et al*. 2009
E. serotinus	Volgograd		ORI	GQ272579	Artyushin *et al*. 2009
E. serotinus	West Kazakhstan		ORI	GQ272574	Artyushin *et al*. 2009
E. serotinus	Astrakhan		ORI	GQ272580	Artyushin *et al*. 2009
E. serotinus	Astrakhan		ORI	GQ272581	Artyushin *et al*. 2009
E. serotinus	Astrakhan		ORI	GQ272586	Artyushin *et al*. 2009
E. serotinus	Astrakhan		ORI	GQ272576	Artyushin *et al*. 2009
E. serotinus	Astrakhan		ORI	GQ272577	Artyushin *et al*. 2009
E. serotinus	Astrakhan		ORI	GQ272574	Artyushin *et al*. 2009

APPENDIX C

The modified Nucleospin® Tissue Protocol for DNA extraction

Standard protocol for human or animal tissue and cultured cells

1. Prepare samples

Cut 25 mg animal tissue into small pieces.

2. Pre-lysis

Add 180 µl buffer T1 and 25 µl proteinase K solution. Vortex to mix. Be sure that the samples are completely covered with lysis solution. Incubate at 56°C until complete lysis is obtained (24h). Use a shaking incubator.

3. Lysis

Vortex samples. Add 200 µl buffer B3, vortex vigorously and incubate at 70°C for 10 minutes. Vortrex briefly.If insoluble particles visible, centrifuge 5 minutes.

4. Adjust DNA binding conditions

Add 210 µl ethanol (96–100 %) to the sample and vortex vigorously.

5. Bind DNA

For each sample, place one NucleoSpin® Tissue column into a 2 ml collecting tube. Apply the same sample to the column. Centrifuge for 1 minute at 11,000 × g. Discard the flow-through and place the column back into the collecting tube.

6. Wash silica membrane

1st wash: Add 500 µl buffer BW. Centrifuge for 1 minute at 11,000× g. Discard flow-through and place the column back into the collecting tube.

2nd wash: Add 600 µl buffer B5 to the column and centrifuge for 1 minute at 11,000 × g. Discard flow-through and place the column back into the collecting tube.

7. Dry silica membrane

Centrifuge the column for 1 minute at 11,000 × g. Residual ethanol is removed during this step.

8. Elute highly pure DNA

Place the column into a 1.5 ml microcentrifuge tube and add 50 µl prewarmed dH2O (70°C). Incubate at room temperature for 1 minute. Centrifuge 1 minute at 13,000 × g.

APPENDIX D

Preparation of the samples for sequencing:

A) Sequencing reaction mix for PCR

dH_20	5.5 µl
Big Dye® Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems)	1.0 µl
Sequencing buffer (included in the kit)	1.5µl

B) Sephadex purification of the sequencing reaction

1. Take Sephadex solution to warm in room temperature 1 hour before starting.

2. When the Sephadex solution has warmed, put it in a magnet mixer.

3. Put a needed amount of Sephadex columns in a rack.

4. Pipette 500 µl of Sephadex solution into each column and let settle for 10 minutes.

5. Centrifuge 4.6 × 1000 rpm for 1 minute.

6. Move the columns into clean Eppendorf tubes.

7. Pipette the sequencing reaction carefully in the middle of the Sephadex column without touching it.

8. Place the Eppendorf tubes with columns in the centrifuge the same way as before (check the position of the column).

9. Centrifuge 4.6 × 1000 rpm for 1 minute.

10. Take the columns from the Eppendorf tubes, discard the gel and rinse the columns with Milli-Q water 4–5 times and put the columns in Milli-Q water into microwave oven to boil for approximately 20 minutes.

11. Place the purified sequencing reactions in a freezer until pipetting into a sequencing plate.

APPENDIX E

Table 1 The 24 variable positions of the 603 bp segment of the cytochrome *b* observed among 92 Northern Bat individuals from this study and from GenBank (see Appendix A for the individuals of each haplotype and sources of the sequences). The transversion is marked with grey background.

Haplotype	Nucleotide position																							
	4	9	64	54	72	80	117	129	141	174	207	249	251	264	276	312	349	359	409	456	471	483	552	597
1	T	C	A	G	C	A	C	C	T	T	T	T	G	C	G	T	G	C	G	T	A	G	T	A
2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.
3	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	C	.	.	.	.	.	.	.	.
4	.	.	.	.	.	G	.	.	.	.	.	.	A	.	.	C	.	.	.	.	.	.	.	.
5	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	A	.	.	.	.	.
6	.	.	G	.	.	.	.	.	.	.	.	.	A	.	.	C	.	.	A	.	.	.	.	G
7	C	.	G	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	A	.	G	.	.	.
8	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	A	.	.	.	.	.
9	.	.	G	.	.	.	T	.	.	.	.	.	.	T	.	C	.	.	A	.	.	.	.	.
10	.	.	G	.	.	.	.	G	.	C	.	C	.	.	.	C	.	.	A	.	.	.	.	.
11	.	.	G	.	T	.	.	.	.	.	.	.	.	.	.	C	.	.	A	.	.	.	.	.
12	.	T	G	.	.	.	.	.	C	.	C	.	A	.	A	C	A	T	A	C	.	A	C	.
13	.	T	G	A	.	.	.	.	C	.	C	.	A	.	A	C	A	T	A	C	.	A	C	.
14	.	T	G	A	.	.	.	.	C	.	C	.	.	.	A	C	A	T	A	C	.	A	C	.

Table 2 The 49 variable positions of the 327 bp HVII sequence observed among 145 Northern Bat individuals by clusters I – V. Individuals belonging to each haplotype are listed in Appendix A. The transversion is marked with grey background.

		Nucleotide position																												
Cluster	**Haplo-type**	1	2	4	24	129	132	165	169	182	184	216	224	225	226	230	236	237	238	239	245	247	249	250	267	269	270	274	276
I	**15**	G	A	T	T	C	A	T	T	C	A	G	C	G	G	A	G	T	A	G	A	G	A	C	A	A	G	T	G
	16	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	17	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	18	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	19	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	20	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.
	21	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	22	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	23	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	24	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	25	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	G	.	.	.
	26	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	27	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	28	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	29	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	30	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.
	31	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.
	32	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.
	33	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.
	34	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.
	35	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	36	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	37	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	38	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Cluster I continues.

Cluster	Haplo-type	Nucleotide position																				
		277	278	279	288	289	290	291	292	297	298	299	301	302	306	307	308	310	313	314	316	327
I	**15**	T	A	G	T	G	A	C	A	A	A	A	T	G	C	G	A	T	A	A	C	G
	16	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.
	17	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.
	18	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.
	19	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	C	.	.	.	.
	20	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.
	21	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	C	.	.	.	.
	22	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.
	23	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.
	24	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.
	25	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.
	26	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	T	.
	27	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	T	A
	28	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	C	.	.	.	.
	29	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	C	.	.	.	.
	30	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	C	.	.	.	.
	31	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	C	.	.	.	.
	32	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	A
	33	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.
	34	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.
	35	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	C	.	.	.	.
	36	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	C	.	.	.	.
	37	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	C	.	.	.	.
	38	.	.	.	.	A	.	.	.	.	G	.	.	.	.	.	.	C	.	.	.	.

Table 2 continues.

Cluster	Haplo-type	Nucleotide position																												
		1	2	4	24	129	132	165	169	182	184	216	224	225	226	230	236	237	238	239	245	247	249	250	267	269	270	274	276
II	39	.	.	.	.	.	.	C	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	40	.	.	C	.	.	.	C	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	41	.	.	.	.	.	.	C	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	42	.	.	.	.	.	.	C	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	43	.	.	.	.	.	.	C	.	.	.	.	.	.	A	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.
	44	.	.	.	.	.	.	C	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A
	45	.	.	.	.	.	.	C	.	.	G	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	46	.	.	.	.	.	.	C	.	.	G	.	.	.	A	.	.	.	.	.	.	.	.	.	G	.	.	.	.
	47	.	.	.	.	.	.	C	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	48	A	.	.	.	.	.	C	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	49	.	.	.	.	.	.	C	.	.	.	.	.	.	A	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	50	.	.	.	.	.	.	C	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
III	51	.	.	.	.	.	.	C	.	.	.	.	T	A	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	52	.	.	.	.	.	.	C	.	.	.	.	T	.	A	.	.	.	.	A	.	.	.	.	G	.	.	.	A
	53	.	.	.	.	.	.	C	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	G	.	.	.	A
	54	.	.	.	.	.	.	C	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	G	.	.	.	A
	55	.	.	.	.	.	.	C	.	.	G	.	.	.	A	.	.	.	G	.	.	.	.	.	G	.	.	.	A
	56	.	.	.	.	.	.	C	.	.	.	.	.	.	A	.	.	.	G	.	.	.	.	.	G	.	.	.	A
	57	.	.	C	.	.	.	C	.	.	.	.	.	.	A	.	.	C	G	.	.	.	.	.	G	.	.	.	A
	58	.	.	C	.	.	.	C	.	.	.	.	.	.	A	.	.	.	G	.	.	.	.	.	G	.	.	.	A

Clusters II and III continue.

Cluster	Haplo-type	Nucleotide position																			
		277	278	279	288	289	290	291	292	297	298	299	301	302	306	307	308	310	313	314	316	327
II	39	C	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	C	.	.	.	.
	40	.	.	.	.	A	.	.	.	G	G	.	.	A	.	.	.	C	.	.	.	.
	41	C	.	.	.	A	.	.	.	G	G	.	.	.	.	.	.	C	.	.	.	.
	42	C	.	.	.	A	.	.	.	.	G	.	.	A	.	.	.	C	.	.	.	.
	43	C	.	.	.	A	.	.	.	.	G	.	.	.	.	.	.	C	.	.	.	.
	44	C	G	.	.	A	.	.	.	.	G	.	.	.	.	.	.	C	.	.	.	A
	45	C	.	.	.	A	.	.	.	.	G	.	.	.	.	.	.	C	.	.	.	.
	46	C	.	.	.	A	.	.	.	.	G	.	C	.	.	.	.	C	.	.	.	.
	47	C	.	.	.	A	.	.	.	.	G	.	.	.	.	.	.	C	.	.	.	.
	48	C	.	.	.	A	.	.	.	.	G	.	.	.	.	.	.	C	.	.	.	.
	49	C	.	.	.	A	.	.	.	.	G	.	.	.	T	.	.	C	.	.	.	.
	50	C	.	.	.	A	.	.	.	.	G	.	.	.	T	.	.	C	.	.	.	.
III	51	.	.	.	.	A	G	.	.	.	G	.	.	.	.	A	.	C	.	.	.	.
	52	.	.	.	.	A	G	.	.	.		.	.	.	.	A	.	C	.	.	.	.
	53	.	.	.	.	A	G	.	G	.	G	.	.	A	.	A	.	C	.	.	.	.
	54	.	.	.	.	A	G	.	.	.	G	.	.	.	.	A	.	C	.	.	.	.
	55	.	.	.	.	.	G	.	.	G	G	.	.	.	.	A	.	C	.	.	.	.
	56	.	.	.	.	A	G	.	.	.	G	.	.	.	.	A	.	C	.	.	.	.
	57	.	G	.	.	A	G	.	.	.	G	.	.	.	.	A	.	C	.	.	.	.
	58	.	G	.	.	A	G	.	.	.	G	.	.	.	.	A	.	C	.	.	.	.

Table 2 continues.

Cluster	Haplo-type	Nucleotide position																												
		1	2	4	24	129	132	165	169	182	184	216	224	225	226	230	236	237	238	239	245	247	249	250	267	269	270	274	276
IV	59	.	.	.	.	.	.	C	.	.	.	A	.	.	A	.	.	.	.	A	.	.	.	T	.	.	A	.	.
	60	.	.	.	.	.	.	C	.	.	.	A	.	.	A	.	.	.	.	A	.	.	.	T	.	.	A	.	.
	61	.	.	.	.	.	.	C	.	.	.	.	.	.	A	.	.	.	.	A	.	.	.	T	.	.	A	C	.
	62	.	.	.	.	.	.	C	.	.	.	.	.	.	A	.	.	.	.	A	.	.	.	T	.	.	A	C	.
	63	.	.	.	.	.	.	C	.	.	.	A	.	.	A	.	.	.	.	A	G	.	.	T	.	.	A	C	.
	64	.	.	.	.	.	.	C	.	.	.	A	.	.	A	.	.	.	.	A	.	.	.	T	.	.	A	C	.
	65	.	.	.	.	T	.	C	.	.	.	A	.	.	A	.	.	.	.	A	.	.	.	T	G	.	A	C	.
	66	.	.	.	.	.	.	C	.	.	.	A	.	.	A	.	.	.	.	A	.	.	.	T	.	.	A	C	.
	67	.	.	.	.	.	.	C	.	.	.	A	.	A	A	.	.	.	.	A	.	.	.	T	.	.	A	C	.
	68	.	.	.	.	.	.	C	.	.	.	A	.	.	A	.	.	.	.	A	.	.	.	T	.	.	A	C	.
	69	.	.	.	.	.	.	C	.	.	G	A	.	.	A	.	.	.	.	A	.	.	.	T	.	.	A	C	.
V	70	.	.	.	.	.	G	C	.	.	.	.	.	.	A	.	.	.	.	A	.	.	.	T	.	.	.	.	.
	71	.	.	C	C	.	G	C	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	T	.	G	.	.	.
	72	A	G	C	C	.	G	C	.	.	.	.	.	.	.	.	.	.	.	A	.	A	G	T	.	G	.	.	.
	73	A	G	C	C	.	G	C	.	.	.	.	.	.	.	.	.	.	.	A	.	A	.	T	.	G	.	.	.
	74	A	.	C	C	.	G	C	.	.	.	.	.	.	.	.	.	.	.	A	.	A	.	T	.	G	.	.	.
	75	A	.	C	C	.	G	C	.	.	.	.	.	.	.	.	.	.	.	A	.	A	.	T	.	G	A	.	.
	76	A	G	C	C	.	G	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	G	.	.	.
	77	A	G	C	C	.	G	C	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	T	.	G	.	.	.
	78	A	G	C	C	.	G	C	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	T	.	G	.	.	.
	79	A	G	C	C	.	G	C	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	T	.	G	.	.	.

Clusters IV and V continue.

Cluster	Haplo-type	Nucleotide position																			
		277	278	279	288	289	290	291	292	297	298	299	301	302	306	307	308	310	313	314	316	327
IV	59	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	C	.	.	.	.
	60	.	.	.	.	A	.	.	.	.	G	.	.	.	.	.	.	C	.	.	.	.
	61	.	.	.	.	A	.	.	.	.	G	.	.	.	.	.	.	C	.	.	.	.
	62	.	.	.	.	A	.	.	.	.	G	.	.	.	.	A	.	C	.	.	.	.
	63	.	.	.	.	A	.	.	.	.	G	.	.	.	.	A	.	C	.	.	.	.
	64	.	.	.	.	A	.	.	.	.	G	.	.	.	.	A	.	C	.	.	.	.
	65	.	.	.	.	A	.	.	.	.	G	.	.	.	.	.	.	C	.	.	.	.
	66	.	.	.	.	A	.	.	.	.	G	.	.	.	.	.	.	C	.	.	.	.
	67	.	.	.	.	A	.	.	.	.	G	.	.	.	.	.	.	C	.	.	.	.
	68	.	.	A	.	A	.	.	.	.	G	.	.	.	.	.	.	C	.	.	.	.
	69	.	.	A	.	A	.	.	.	.	G	.	.	.	.	.	.	C	.	.	.	.
V	70	.	.	.	.	.	.	.	.	.	G	.	.	.	.	A	.	C	.	.	.	.
	71	.	G	.	G	.	.	A	G	G	.	G	.	A	.	A	.	C	G	.	.	.
	72	.	G	.	G	.	.	A	G	G	.	G	.	A	.	A	.	C	.	.	.	.
	73	.	G	.	G	.	.	A	G	G	.	G	.	A	.	A	G	C	.	.	.	.
	74	.	G	.	G	.	.	A	G	G	.	G	.	A	.	A	.	C	.	.	.	.
	75	.	G	.	G	.	.	A	G	G	.	G	.	A	.	A	.	C	.	.	.	.
	76	.	G	.	G	.	.	A	G	G	.	G	.	A	.	A	.	C	.	G	.	.
	77	.	G	.	G	T	.	A	G	G	.	G	.	A	.	A	.	C	.	.	.	.
	78	.	G	.	G	.	.	A	G	G	.	G	.	A	.	A	.	C	.	.	.	.
	79	.	G	.	G	.	.	A	G	G	.	G	.	A	.	A	.	C	G	.	.	.